DOCUMENT RESUME

ED 375 312 CE 067 480

TITLE	Alternative Designs for Evaluating Workplace Literacy Programs. Conference Proceedings and Commissioned Papers at the "Design Guidance for Evaluating Workplace Literacy Programs" Work Group Conference (Washington, D.C. April 13, 1993).
INSTITUTION	Research Triangle Inst., Research Triangle Park, N.C.
SPONS AGENCY	Department of Education, Washington, DC. Office of the Under Secretary.
REPORT NO	ED/OUS-94-18
PUB DATE	Jun 94
CONTRACT	LC92-008-001
NOTE	109p.
PUB TYPE	Collected Works - Conference Proceedings (021)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Adult Basic Education; Adult Programs; Articulation (Education); *Educational Research; Evaluation Criteria; *Evaluation Methods; *Evaluation Research; *Literacy Education; Longitudinal Studies; Models; Outcomes of Education; *Program Evaluation; *Research Methodology
IDENTIFIERS	*Workplace Literacy

ABSTRACT

This document contains the five papers presented at a meeting at which key issues in evaluating workplace literacy programs were discussed. In "Key Components of Workplace Liteacy Projects and Definitions of Project 'Modules,'" Judith A. Alamprese describes the context for evaluating the National Extension Program, components of workplace literacy programs, and validation of program components. Design alternatives for a national program evaluation and design options for an effective practices study are explored by Jorie W. Philippi in "Measuring Workplace Literacy Program Effectiveness: Evaluation Purposes and Methodologies." In "Advantages and Disadvantages of Longitudinal Designs for Evaluating Workplace Literacy Programs," Elisabeth Hayes examines the following: the short-term nature of literacy gains, random assignment versus other methods, availability of comparison/control groups, and practical feasibility of alternative designs and procedures. The last two papers, "Articulation and Measurement of Participant Outcomes in Workplace Literacy Programs" (Larry Mikulecky) and "Articulation and Measurement of Program Outcomes" (Anthony R. Sarmiento), focus on the relationships between curriculum/instructional design and various participant/program outcomes and assess the availability and appropriateness of data and various data collection methods. Most papers include substantial bibliographies. Appended are the work group conference agenda, biographies of the presenters, and a list of work group participants. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 375 312

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CE067480

BEST COPY AVAILABLE

The views expressed in this report, developed under contract with the U.S. Department of Education, do not necessarily reflect the position or policy of the Department, and no official endorsement by the Department should be inferred.

BEST COPY AVAILABLE

ALTERNATIVE DESIGNS FOR EVALUATING WORKPLACE LITERACY PROGRAMS

Conference Proceedings and Commissioned Papers

March 1994

Prepared for:

U.S. Department of Education
Office of Policy and Planning

Prepared by:

MSPD Evaluation Support Center
Research Triangle Institute
P.O. Box 12194
Research Triangle Park, NC 27709

Contract No. LC92-008-001

Contents

Introduction i

Chapter 1 — Key Components of Workplace Literacy Projects and Definition of Project "Models"
Author: Judith A. Alamprese 1-1

Chapter 2 — Measuring Workplace Literacy Program Effectiveness: Evaluation Purposes and Methodologies
Author: Jorie W. Philippi 2-1

Chapter 3 — Advantages and Disadvantages of Longitudinal Designs for Evaluating Workplace Literacy Programs
Author: Elisabeth Hayes 3-1

Chapter 4 — Articulation and Measurement of Participant Outcomes In Workplace Literacy Programs
Author: Larry Mikulecky 4-1

Chapter 5 — Articulation and Measurement of Program Outcomes
Author: Anthony R. Sarmiento 5-1

Appendix A — Work Group Conference Agenda
Appendix B — Biographies of Presenters
Appendix C — Work Group Participants

INTRODUCTION

In response to continuing concerns that the nation's workforce lacks the basic skills to ensure our competitiveness in the global economy, Congress created the National Workplace Literacy Program (NWLP) in 1988. Originally contained in the Omnibus Trade and Competitiveness Act of 1988 (P.L.100-418), the program was incorporated in the 1988 Hawkins-Stafford Amendments (P.L. 100-297) to facilitate speedy implementation. It was subsequently amended by the National Literacy Act of 1991 (P.L. 102-73). The major purpose of NWLP is to support local projects established as partnerships between business/industry/labor and education that provide services ". . .to improve the productivity of the workforce through improvement of literacy skills needed in the workplace" (P.L. 102-73, § 371(a)(3)). Grantees are authorized to address this purpose through a variety of activities, including:

(A) providing adult literacy and other basic skills services and activities;

(B) providing adult secondary education services and activities which may lead to the completion of a high school diploma or its equivalent;

(C) meeting the literacy needs of adults with limited English proficiency;

(D) upgrading or updating basic skills of adult workers in accordance with changes in workplace requirements, technology, products, or processes;

(E) improving the competency of adult workers in speaking, listening, reasoning, and problem solving; or

(F) providing education counseling, transportation, and nonworking hours child care services to adult workers while they participate in a program funded under paragraph (2) (A) (P.L. 102-73, §371).

Administered by the Division of Adult Education and Literacy (DAEL) and the Division of National Programs (DNP) in the Education Department's Office of Vocational and Adult Education (OVAE), NWLP completed five funding cycles through FY 1992. In these funding cycles, ED awarded 261 grants to projects in 45 states and territories for a total of $82 million. The average

grant size over the first five funding cycles was $313,242. Overall, these 261 grants provided services to nearly 120,000 workers (U.S. Department of Education, 1993).

The importance of the services funded under NWLP is widely acknowledged, and Congress and ED have made ongoing adjustments to improve the efficiency and facilitate the effectiveness of projects funded by the program. Even so, program proponents, administrators, and researchers have many questions regarding the types of activities and services that may be most effective in supporting the goals of NWLP. Responding to the need for research-based information to gauge the effectiveness of the myriad project goals, instructional approaches, and service configurations implemented at the local level throughout the program's relatively brief history, the Department of Education's Planning and Evaluation Service (PES) and DAEL designed an evaluation of effective workplace literacy programs for initiation in fall 1993. To inform the planning for the evaluation, PES awarded a Task Order to Research Triangle Institute, the activities of which included identification of topics for investigation, establishment and convening of a work group to consider key topics, arrangement for preparation of papers by experts on key topics in workplace literacy, and preparation of a report containing the papers. ED officials in PES and DAEL guided all phases of the effort, the key activities of which included the following.

Selection of Topics for the Work Group Presentations and Papers

A key purpose of the Task Order was to examine issues central to design and conduct of a national study that would address topics of interest to ED, particularly determination of "what works" in workplace literacy projects. Review of available research and discussions with PES and DAEL staff permitted selection of five topics to be addressed in the Work Group meeting and subsequent issue papers. These topics, and the experts who agreed to make presentations and prepare papers for inclusion in this volume, are as follows.[1]

[1]The meeting agenda is presented in Appendix A; brief descriptions of presenters' professional experience are contained in Appendix B.

Key components of workplace literacy projects and definition of project "models"

- Partnerships, institutionalization strategies, dissemination, contextual learning activities, focused job task analyses, employee involvement in planning, provision of support services, etc.
- Articulation of 5 to 10 project "models" reflecting salient combinations of components

Presenter: Judy Alamprese

Evaluation purposes and methodological alternatives

- Design alternatives for a national program evaluation
- Design options for an effective practices study
- Determination of key evaluation questions to address identified study purposes:
  - Effectiveness of workplace literacy projects
  - Identification of promising practices
  - National Workplace Literacy Program (NWLP) accomplishments
- Requirements for incorporating national and project-level analyses in a national study

Presenter: Jorie Philippi

Advantages and disadvantages of a longitudinal design

- Short-term nature of literacy gains
- Locating study participants over time
- Random assignment versus other methods
- Availability of comparison/control groups
- Influence of small treatment group size and attrition problems on analytic power
- Practical feasibility of alternative designs and procedures

Presenter: Elisabeth Hayes

Articulation and measurement of program outcomes

- Improved basic skills, critical thinking skills, productivity, safety, career advancement, literacy self-image; earnings gains; increased literacy activity; GED; English-language proficiency; other
- Relationships among curriculum and instructional design, project and participant goals, and participant outcomes
- Availability and appropriateness of tests and other data collection methods

Presenter: Larry Mikulecky

Articulation and measurement of program outcomes

- Changes in organizational culture, increased productivity, safety, worker empowerment, supervisor satisfaction, other
- Availability and utility of business data
- Project institutionalization, dissemination of project models or materials, replication or adaptation, other

Presenter: Tony Sarmiento

Summary of key issues in workplace literacy research and evaluation

Presenter: Tom Sticht

Work Group Meeting on Design Alternatives

The Work Group meeting, held in Washington, DC, on April 13, 1993, provided the opportunity for persons with a range of interests and expertise to discuss key issues in evaluating workplace literacy programs. Persons attending included directors of projects funded under NWLP, federal policymakers, university-based and other researchers, and persons involved in providing

technical assistance and other services to local NWLP-funded projects.[2] The meeting was organized around the six presentations, with discussions among participants based on each of the topics following the presentation. After the meeting, PES made available an audiotape of the proceedings.

Organization of the Report

Subsequent sections of this document contain the papers prepared by each of the experts who made presentations at the Work Group meeting. Each author received a summary of the Work Group discussion for use in developing the final paper, and each of the papers reflects those discussions and additional information the author chose to incorporate in the paper. Papers appear in the order of presentations at the conference.

[2]Appendix C provides a list of participants in the Work Group meeting.

KEY COMPONENTS OF WORKPLACE LITERACY PROJECTS AND DEFINITION OF PROJECT "MODELS"

Judith A. Alamprese
COSMOS Corporation

Abstract

The U.S. Department of Education's National Workplace Literacy Program (NWLP) provides grants for demonstration projects that are designed to improve the productivity of the workforce through the improvement of literacy skills needed in the workplace. As a demonstration program, one intent of the NWLP is to identify effective models of workplace literacy practice or the components that constitute effective models. While little empirical data exist concerning the types of workplace literacy program services that result in organizational productivity and learner skill enhancements, there is information about the types of services that are being carried out in NWLP-funded projects and their perceived utility. This issue paper reviews the state-of-knowledge concerning the key components and "models" of workplace literacy programs, and describes methodologies that can be used to evaluate these components and models.

Introduction

Authorized under the Adult Education Act (Public Law 100-297, as amended by the National Literacy Act of 1991, Public Law 102-73), the U.S. Department of Education's National Workplace Literacy Program (NWLP) provides grants for demonstration projects that are designed to improve the productivity of the workforce through the improvement of literacy skills needed in the workplace. Since fiscal year 1988, the NWLP has funded five cycles of grantees for the purpose of developing partnerships to carry out workplace literacy programs.

One intent of the legislation authorizing the NWLP is to identify effective models of workplace literacy practice or, at least, the components of workplace programs that constitute effective models. With the exception of the first cycle of grantees, each NWLP project was to conduct an independent evaluation to determine the strengths and weaknesses of the services provided as well as the client and organizational outcomes resulting from these services. In addition to these project-level evaluations, the legislation calls for an evaluation of the overall NWLP.

As part of the planning for a national evaluation of the NWLP, the U.S. Department of Education commissioned adult literacy researchers to prepare papers that addressed the key issues involved in designing a national evaluation. This issue paper discusses the state-of-knowledge concerning the key components and "models" of workplace literacy projects, and describes methodologies that can be used to evaluate these components and models.

Context for Evaluating the NWLP

Any development of a framework for evaluating the NWLP should consider the role of the NWLP as a national demonstration program. Since the 1960s, the federal government has used demonstration projects--action programs that operate in a real-life setting--to develop models of services for addressing social problems in education, job training, human services, and other areas. Work by the Rand Corporation (Baer, Johnson, & Merrow, 1976; Glennan et

al., 1978) has helped to define the purposes of demonstrations as either policy-implementing or policy-formulating. Policy-implementing demonstrations test ideas by taking research findings and using them in everyday practices. In a policy-formulating demonstration, new, field-based ideas are produced that can then be analyzed by researchers under controlled settings. Both types of demonstrations are expected to produce information about the implementation of a program in a real-life setting.

In recent years, the U.S. Department of Education has funded evaluations of demonstration programs to identify different models for implementing services. In the areas of vocational and adult education, demonstration programs such as the Cooperative Demonstration Program (High Technology) have been assessed to determine the utility of the partnerships formed by school districts, community colleges, consortia, and private industry to demonstrate new approaches to vocational education (Bateman, Muraskin, Adduci, & Boesel, 1992). While these evaluations have been useful in describing the types of services provided by demonstration programs and the conditions under which these services are thought to be effective, there has been an increasing emphasis by the Department of Education on linking program operations with outcomes so that the most critical program components can be identified and replicated (Ginsburg, 1992).

A national evaluation of the NWLP demonstration program provides a distinctive opportunity for adult education. Little empirical data exist concerning the effectiveness of the types of services that have been provided by projects funded under the NWLP or workplace literacy programs that have been sponsored by other public or private entities. Rather, the assessments that have been undertaken offer a preliminary understanding of the program services that have been implemented and their potential for enhancing workers' literacy skills. Through a national evaluation, the overall utility of the NWLP demonstration program can be assessed in terms of the extent to which it has been either policy-implementing or policy-formulating in meeting the objectives of the legislation. Furthermore, systematic data can be collected about the partnership activities, staff training, and instructional services that have been carried out and the impacts of these activities on learners' skill improvement and the productivity of the workplaces in which these activities are taking place.

Components of Workplace Literacy Programs

While there is limited empirical evidence about the utility of the services carried out in workplace literacy programs, two types of information do exist that can be used in developing an evaluation design to identify and validate the components of workplace literacy programs. The first is the regulations for the NWLP, which provide guidance about the types of partnership members and services that can be included in a workplace literacy demonstration project. The second is the literature on workplace literacy programs, which consists of NWLP project reviews, guidebooks for developing workplace literacy programs, frameworks for evaluating programs, and reports from project-level evaluations of NWLP grantees.

Guidance from NWLP Regulations

The NWLP grants announcements published in the *Federal Register* have specified the types of organizational entities that can be involved in a program and the kinds of services that are eligible for funding. A unique aspect of the NWLP is the partnership requirement. The types of partners that must participate in a NWLP project are the following:

a) A business, industry, or labor organization, or private industry council; and

b) A state educational agency, local educational agency, institution of higher education, or school (including an area vocational school, an employment and training agency, or a community-based organization (*Federal Register*, June 5, 1992, p. 24130)).

The regulations stipulate that a partnership must include at least one entity from (a) and at least one entity from (b), and may include more than one entity from each group.

Also described are the types of services that can be carried out by a NWLP project. The following six activities are allowed:

- Adult literacy and other basic skills and services;
- Adult secondary education services and activities that may lead to the completion of a high school diploma or its equivalent;
- Services for adults with limited English proficiency;

- Services for upgrading or updating basic skills of adult workers in accordance with changes in workplace requirements, technology, products, or processes;

- Services for improving the competency of adult workers in speaking, listening, reasoning, and problem solving; or

- Educational counseling, transportation, and child care services for adult workers during nonworking hours while workers participate in the services offered by the project.

In addition to providing instructional services, a NWLP project is supposed to assess the progress of participants, evaluate the project's impact, and disseminate information about the project.

This information has been used by prospective grantees in formulating their project plans and can be considered as one source of data about the types of components of service that a NWLP project should comprise.

State of Knowledge Concerning Program Components

As in any new field, the development of a knowledge and research base about workplace literacy practice has been gradual. Most studies have been descriptive rather than analytical, and the available documentation has included overviews of program activities and frameworks for developing and evaluating workplace literacy programs. In spite of this lack of empirical information, the existing reports and guidebooks do provide knowledge that can be used to specify hypotheses about effective program models and components. For example, reviews of workplace literacy programs have identified the key components reported to be associated with effective workplace literacy programs and methods for improving program effectiveness (e.g., Cornell, 1988; Division of Adult Education and Literacy, 1992; Kutner et al., 1991; Stein, 1991). Information about workplace literacy program components also is presented in guidebooks, which delineate the steps that should be taken to design and implement a workplace literacy program (e.g., Henard, Lloyd, & Mikulecky, 1992; Sticht 1991a). While often not specifically citing program components, reports that discuss methods and issues related to the evaluation of workplace literacy program offer information about the key features of an effective program (e.g., Philippi, 1992; Sticht, 1991b; Sticht, 1992). Finally, evaluations of NWLP-funded projects and other workplace literacy programs provide

insights about the components of programs that appear to have been successful and those that have been problematic in implementation (e.g., Alamprese & Kay, 1993; Mikulecky & Lloyd, 1992; Program Planning Consultant, 1992).

A synthesis of the information provided in these studies and reports reveals a common set of components that seem to be important for carrying out an effective workplace literacy program. While similar components and corresponding activities are cited in this literature, there is little quantification or measurement of the extent to which activities should be carried out--a key aspect of model building. The key components discussed are the following:

- Partnership;
- Staff selection and training;
- Operational plan;
- Curriculum and instruction;
- Support services;
- Learner assessment;
- Program evaluation; and
- Dissemination.

In the literature, various characteristics of practice and activities that constitute each component are presented. While it is not assumed that a program must implement all of these activities in order to be effective, this information does provides a starting point from which to hypothesize the types of activities that may be linked with program outcomes.

Partnership. In workplace literacy partnerships, the literature suggests that effective practice includes the following:

- The partners agree on program goals, identify benefits, and have a set of common expectations about the activities that will be carried out by the project.
- There is a provision for shared governance (e.g., an advisory committee with representation from participating workers, management, labor, and instructional staff).

- There is ongoing involvement by the partners (e.g., providing classroom space, assisting in the monitoring of project services).
- There is support from upper management and on-line supervisors.
- There is a prior working relationship among the partners.
- There are incentives for workers to participate that are usually stated in the partnership agreement (e.g., release time, training bonuses).

Staff selection and training. Staff should have some background in teaching adults (e.g., adult basic education, English as a second language), and preferably experience with workplace literacy programs. Given the likelihood that few staff will have had prior experience with workplace literacy since it is a developing field, the provision of inservice training appears to be critical to the success of a project. Ongoing training should be offered that emphasizes the unique and common aspects of workplace literacy program operations, the factors associated with the implementation of a demonstration program, and the uses of data and information in monitoring project activities.

Operational plan. A conceptual framework should be developed for each project that includes:

- Assumptions guiding the project, and activities that are expected to produce outcomes; and
- The population of workers targeted for the services that are to be offered and the recruitment strategies that will be used to attract the participation of these workers.

Curriculum and instruction. Of all of the program components, the area of curriculum is discussed the most comprehensively in the literature. There is general agreement that some type of job analysis should be conducted that involves the workers for whom services are being designed. Based on the results of the task analysis, curriculum should be developed as a conceptual system with objectives and job-based competencies that comprise underlying skills which are transferrable to other settings. The curriculum should emphasize contextual learning and incorporate actual materials from the workplace. The use of these materials facilitates the teaching of skills in context and heightens learners' interest.

The development of curriculum also should involve workers and focus on teaching literacy skills that mediate job performance.

Many workplace literacy programs have prepared individualized learning plans for workers to reinforce their learning goals and focus instruction. Often supplemental tutoring is provided to assist workers with developing the skills that they find difficult. Another strategy for facilitating workers' learning is to use multiple learning tools, such as videos, instructional software, and calculators. The use of these tools takes into account workers' different learning styles and learning needs.

The aspect of instruction that is the subject of much discussion is the amount of time that instruction is offered. The quantity and intensity of instruction should match the program's objectives, and information should be collected to determine the minimum amount of instructional time that workers need to enhance their skills.

Support services. In designing a workplace literacy program, consideration should be given to workers' needs for support services. Two types of services frequently provided are child care and transportation. The provision of educational counseling also may help to motivate workers and to encourage their ongoing participation in educational programs.

Learner assessment. A critical component of any workplace literacy program is learner assessment. In keeping with one of the principal tenets of effective instruction, periodic feedback should be given to workers to motivate them and guide their learning activities. Furthermore, the collection of assessment data provides information about the overall performance of a program.

One criterion for selecting assessment instruments is that they reflect the content of instruction. Both standardized tests and applied performance assessments are appropriate for use in workplace literacy programs. Applied performance assessments are particularly effective because they can be used to assess basic skills in the specific context of the workplace. The assessment component of a program should include the administration of placement, pretests, and posttests. The use of these instruments helps to assure that workers will be matched with the most appropriate level of instruction and that their overall learning will be documented. To monitor the ongoing progress of workers, instructors should collect information using methods such as observations, materials review, and informal discussions.

Program evaluation. Both formative and summative evaluation activities should be carried out in a demonstration project to collect the information needed in model building. An initial evaluation plan should be prepared that includes the overall design and data collection and analysis procedures. It is important that this design relate to the project's goals and activities to help assure that the information collected during the evaluation will be utilized.

Dissemination. An important program component that often is neglected in projects is dissemination. Workplace literacy projects should identify their potential audiences early in the project and determine appropriate events for reaching these audiences. Furthermore, the types of materials and information that can be disseminated need to be determined, and a plan should be developed for distributing information to the targeted audiences at the appropriate events. By carrying out dissemination activities, a workplace literacy program can enhance both its visibility in the workplace and the support received from project partners.

Institutionalization. While discussed in the NWLP regulations as a desired outcome from the demonstration effort, not much is known about the institutionalization of workplace literacy programs from the available literature. One reason is that individual projects are having difficulty convincing employers about the value of the long-term sponsorship of the program. Additional efforts should be carried out to assist projects in sustaining their services after the termination of the grant.

Summary

This discussion on workplace literacy program components provides a framework for understanding the types of activities that adult literacy researchers and practitioners have found to be critical to the success of a program. However, much of this guidance is imprecise and does not address specific activity levels. Further information is needed about the extent to which these activities should be carried out, and the relationship between levels of activities and achievement of program outcomes.

Validating Program Components

One methodology that can be used in a national evaluation of the NWLP to validate the critical components that a workplace literacy program should comprise is the case study method. This method, as defined by Yin (1984/1989, p.23), is an empirical inquiry that

> *investigates a contemporary phenomenon within its real-life context, addresses a situation in which the boundaries between phenomenon and context are not clearly evident, and uses multiple sources of evidence.*

Two aspects of the case study method make it particularly appropriate for use in evaluating the NWLP. First, the context of the workplace can be examined in investigating both the intervention (i.e., the workplace literacy project) and the process used in implementing the intervention. Second, the case study is not limited to either qualitative or quantitative data, but can incorporate both types of evidence.

While the case study method often has been used to develop new hypotheses about phenomena, it also can meet the evaluation needs of assessing outcomes and testing hypotheses. For the evaluation of the NWLP, this aspect of the case study method is particularly appealing.

Described below is an overview of the major steps that should be undertaken in using the case study method to evaluate the components of a workplace literacy program (see Yin, 1993, for a detailed discussion of these steps).

Develop a hypothesized understanding of the program being evaluated.

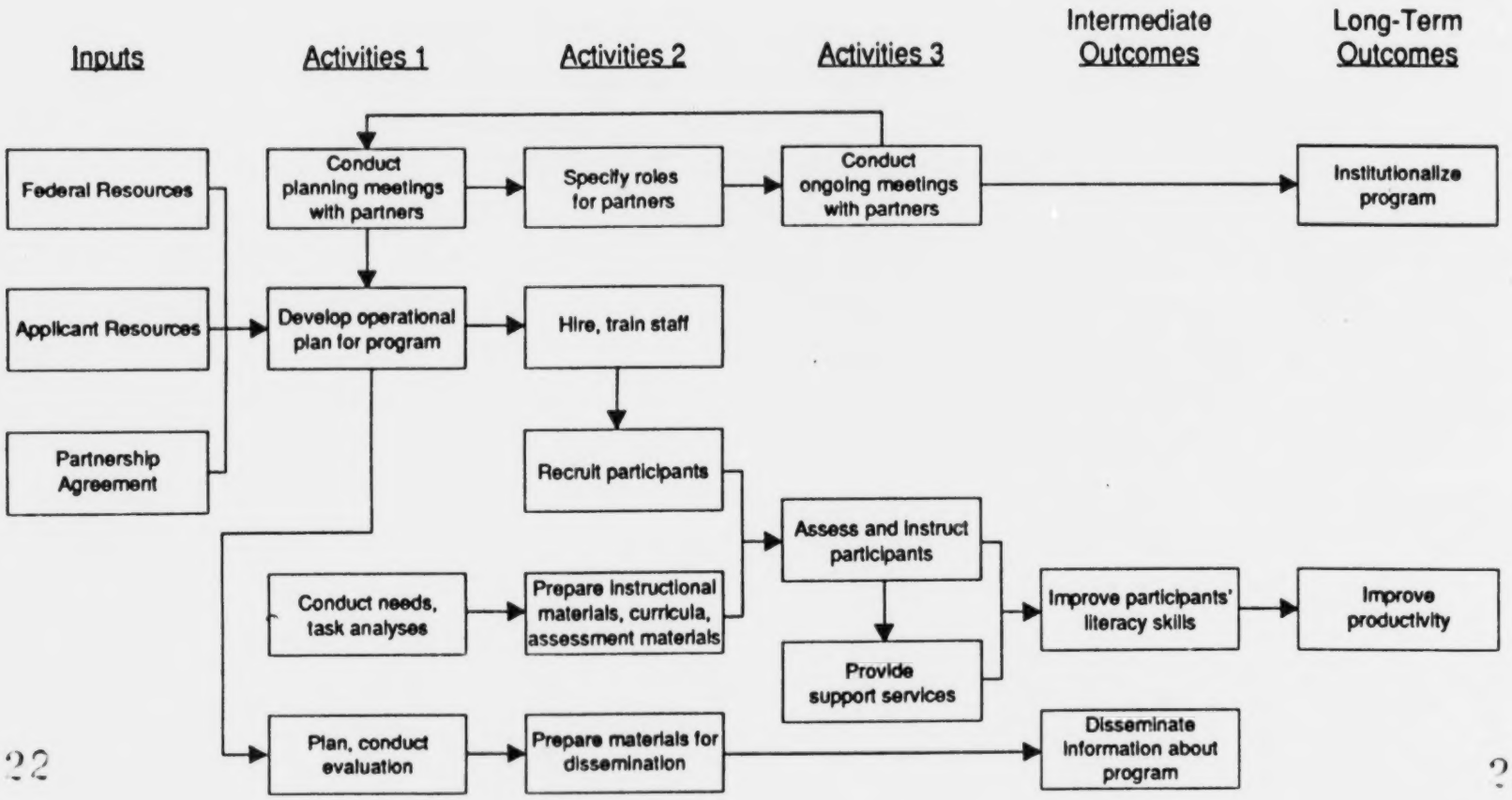
The initial design of the case study evaluation requires an understanding of the program's intended operations and outcomes, with a particular focus on the contextual conditions. An understanding of the program can be reflected in two ways: a program logic model tracing the causal flows of the program (Wholey, 1979), and an emerging taxonomy of contextual conditions within which the program operates.

As an example of how the NWLP can be understood, presented in Figure 1 is an illustrative logic model of a workplace literacy project funded by the NWLP. The model shows the components of a program as specified in the regulations, and the presumed linkages among them. Inputs to the system include federal resources, applicant match and resources,

Figure 1

NATIONAL WORKPLACE LITERACY PROGRAM

ILLUSTRATIVE LOGIC MODEL



BEST COPY AVAILABLE

and the partnership agreement that lays out the conditions under which the members will interact. These inputs should lead to the development of an operational plan, which in turn leads to a number of activities. These include:

- conducting needs and task analyses;
- hiring and training staff;
- determining how the partnership members will work together, and what their respective roles will be;
- recruiting participants;
- preparing instructional materials, curricula, and assessment materials;
- carrying out assessment and evaluation activities; and
- preparing materials for dissemination.

The portrayal of the interrelationships among the activities in the figure is one conception of how a program can operate. Any given project may involve more or fewer activities, and these activities may relate differently than is illustrated in the model. In Activity 3 on the model, the ongoing flow of program operations is shown including partner meetings and the provision of instructional, assessment, and support services. These activities are expected to achieve two types of outcomes:

- intermediate outcomes--improving participants' literacy skills and the conduct of dissemination activities; and
- long-term outcomes--improving productivity in the workplace and institutionalizing the program.

This logic model is a useful tool for determining the types of components and activities that constitute a workplace literacy program, and the ways in which these activities are supposed to result in outcomes.

Immerse this understanding within previous research.

The hypothesized understanding of the program should be placed in the broader context of what is known from research and practice. If possible, rival theories and

hypotheses should be specified. For example, the previous discussion concerning the literature on program components provides background for understanding the portrayal of a workplace program in the logic model and may help to clarify the types of activities that have been included in the model. The broader realm of theory and practice will be used to generalize the results of the evaluation.

Tentatively define the main and subordinate units of analysis.

For the purposes of the NWLP evaluation, the federal program is the main unit of analysis. The subordinate units of analysis include the individual projects funded through the NWLP and the workers who participate in these projects.

Establish a schedule and procedure for making interim and final reports.

An important characteristic of a case study is its ability to provide information throughout the life cycle of an evaluation. To maximize this feature, clear schedules and procedures for submitting information and conducting briefings with program and sponsoring officials should be established.

Define and test instruments, protocols, and field procedures.

Different instruments and procedures are needed for each unit of analysis. The case study method can use any relevant data collection procedure, including fieldwork and participant-observation, surveys, and document analysis. The variables and categories of data that are to be collected should reflect the hypothesized understanding of the program and any rival theories that have been specified. Before the final data collection plans and procedures are determined, a pilot test should be conducted to determine the appropriateness of the instruments and the types of evidence that are likely to be available. For example, the first cycle of the NWLP did not emphasize the conduct of an evaluation. Therefore, the types of data that are likely to be available from these projects will be different from those available from grantees funded in subsequent cycles.

Collect, analyze, and synthesize data.

Often in a case study, data collection and analysis are likely to occur in an intermingled form. Since the sources of evidence vary from case to case, the adequacy of the evidence must be evaluated as it is collected. This process requires an experienced investigator who can identify relevant evidence even though the sources may vary, and can document the methodological steps that have been taken to assure an unbiased data collection process.

Create a case study database.

After data collection is completed, a formal database is created that includes all of the qualitative and quantitative data that were collected. The systematic archiving of this information will facilitate the analysis and reporting processes that will be carried out.

Analyze the evidence.

The analytic techniques used in a case study depend on the types of evidence that were gathered. For example, an evaluation of the NWLP might include qualitative analyses of partnership and implementation processes carried out by projects as well as quantitative analyses of learning gains achieved by individuals participating in the projects.

Compose the case study report.

The final step is to develop a report that is separate from the database and that includes explicit presentations of the evidence used in drawing conclusions.

Conclusion

In describing the key components of a workplace literacy program identified in the literature as critical to the delivery of effective services to clients, this paper has provided a framework for validating program models. An important aspect of this validation process is the development of a logic model tracing the casual flows of the program and a taxonomy of the contextual conditions within which the program operates. As illustrated in the paper, the case study method is a useful tool for evaluating the effectiveness of program components

and for identifying the conditions under which these components can function as a program model.

Reference List

Alamprese, J., & Kay, A. (1993). *Literacy on the cafeteria line: Evaluation of the skills enhancement training program.* Washington: COSMOS Corporation.

Baer, W. S., Johnson, L., & Merrow, E. W. (1976). *Analysis of federally funded demonstration projects.* Santa Monica: The Rand Corporation.

Bateman, P., Muraskin, L., Adduci, L., & Boesel, J. (1992). *Demonstrating cooperation: Lessons from federally-funded projects in vocational education.* Washington: COSMOS Corporation.

Cornell, T. (1988, April). Characteristics of effective occupational literacy programs. *Journal of Reading*, 31(7), 654-656.

Division of Adult Education and Literacy (1992). *Workplace literacy: Reshaping the American workforce.* Washington: U.S. Department of Education.

Ginsburg, A. (1992, Fall). Meeting the Market for Quality Evaluation in Education. In C. W. Wye & R. C. Sonnichsen (Eds.), *Evaluation in the federal government: Changes, trends, and opportunities*, 55, (pp. 37-46). San Francisco: Jossey-Bass Publishers.

Glennan, T. K. et al. (1978, May). *The role of demonstrations in federal r & d policy.* Santa Monica: The Rand Corporation.

Henard, D., Lloyd, P., & Mikulecky, L. (1992). *A guidebook for developing workplace literacy programs.* Bloomington: Indiana University.

Kutner, M., Sherman, R., Webb, L., & Fisher, C. (1991). *A review of the national workplace literacy program.* Washington: Pelavin Associates, Inc.

Mikulecky, L., & Lloyd, P. (1992). *Evaluating the impact of workplace literacy programs.* Bloomington: Indiana University.

Philippi, J. (1992). *How do you know its working? Evaluating the effectiveness of workplace literacy programs.* Springfield: Performance Plus Learning Consultants, Inc.

Program Planning Consultant Group (1992). *Project rise: Workplace litera · education in context.* New Brunswick: Rutgers University.

Stein, S. G. (1991, October). *Tradition and change: The role of workplace education in the transformation of the workplace.* Paper presented at the annual meeting of the American Association for Adult and Continuing Education, Montreal.

Sticht, T. G. (1991a). *Knowledge resources for designing and delivering job-linked literacy programs.* Paper prepared for Work in America Institute, New York.

Sticht, T. G. (1991b). *Evaluating national workplace literacy programs.* El Cajon: Applied Behavioral & Cognitive Sciences, Inc.

Sticht, T. G. (1992). *Evaluation of job-linked literacy programs. Paper prepared for Work in America Institute*, New York.

Wholey, J. (1979). *Evaluation: Promise and performance.* Washington: The Urban Institute. Yin, R. K. (1993). *Applications of case study research.* Newbury Park: SAGE Publications, Inc.

MEASURING WORKPLACE LITERACY PROGRAM EFFECTIVENESS: EVALUATION PURPOSES AND METHODOLOGIES

Jorie W. Philippi
Performance Plus Learning Consultants, Inc.

Abstract

Workplace literacy programs stem from the desire for joint survival by both employees and the organizations for which they work. For this reason, in the workplace, literacy programs are viewed as components of "training," and are usually conducted under the direction of personnel or human resource development departments. As a function of training, workplace literacy programs need to be evaluated differently from other adult education programs. Any effective training program, e.g., a workplace literacy program, demonstrates positive impact both on organizational needs and on those of individual participants. Properly carried out, the process of program evaluation should provide information that improves the effectiveness of program design, development, implementation, and operation. It should also provide hard data supports for identification of indicators of program effectiveness, measurement of change in terms of those indicators, and work toward establishing program standards from the results. Further, evaluation should act as a vehicle for gathering information from many programs in order to draw conclusions about best practices and add to the existing body of knowledge in this emerging field.

Evaluation may be conducted for a variety of reasons that include provision of feedback to program decisionmakers for program quality control; comparisons of techniques and variables to control program criteria; research to prove or disprove specific hypotheses; intervention in an organization to promote change; and exertion of political power within an organization. Identifying the purposes for evaluation and selecting a model and methodology to achieve those purposes are critical to obtaining the desired information. Selecting local program indicators and developing data collection instruments and procedures for measuring changes in those indicators are the core of any good workplace literacy training program evaluation. Indicators should be workplace specific. Analysis, interpretation, and reporting of data obtained should be conducted in ways that are meaningful to program decisionmakers at all levels.

Introduction

Workplace literacy programs stem from the desire for joint survival by both employees and the organizations for which they work (Fields, 1986). Today's economy and workplace environment are changing rapidly, accelerating the need for workers to use literacy skills as they perform job tasks. The shifts toward self-directed teamwork and emphasis on quality production and services necessary for competing in a global marketplace have created job tasks that employ numerous applications of literacy skills. Both large and small companies are requiring many new job tasks of their employees in efforts to remain in business and to leapfrog their competitors.

Because of the necessity for businesses to remain aggressive players in a global marketplace during the reskilling of their labor forces, the workplace has become a vast learning community with immediate and critical needs. To accomplish new workplace tasks effectively and to improve current job performance, employers have begun to realize the necessity for providing workers with training in job literacy skills. Most employees also feel the need for, and often ask for, a "brush up" or special training to prepare themselves to assume control of their work environment. Even those workplace literacy programs now sponsored by labor organizations have expanded beyond adult basic education courses and tuition reimbursement plans as a direct result of member requests for augmentation of company training in new job requirements (Philippi, 1992). For this reason, in the workplace, literacy programs are viewed as components of "training," and are usually conducted under the direction of personnel or human resource development departments. As a function of training, workplace literacy programs need to be evaluated differently from other adult education programs. "Training," by definition, is provided in response to organizational needs as well as those of the individual; conversely, education is a response to individual needs alone. Any effective training program (e.g., a workplace literacy program) demonstrates positive impact both on organizational needs *and* on those of individual participants.

Evaluation is generally thought to be the concluding step in a series of program components, or the completion of a training cycle. It is often described as a summation in terms of "outcomes," and is frequently viewed as a somewhat frivolous program afterthought or a small but necessary requirement fulfilled to remain in compliance with regulations

imposed by an external funding source. In reality, evaluation is an integral part of any workplace literacy training program design; it dictates and monitors data collection to determine summative cyclical outcomes as they relate to an ongoing formative program process.

Properly carried out, the process of program evaluation should provide information that improves the effectiveness of program design, development, implementation, and operation (Sticht, 1991). It should also provide hard data that facilitates identification of indicators of program effectiveness, measurement of change in terms of those indicators, and work toward establishing program standards from the results (Condelli, 1992). And, evaluation should serve as a vehicle for gathering information from many programs in order to draw conclusions about best practices and add to the existing body of knowledge in this emerging field (Philippi, 1993).

Purposes for Program Evaluation

Evaluation can serve a number of purposes. Five main categories are frequently mentioned in the available literature on this topic:

1. Evaluation can be a means *for providing feedback* that enables quality control over the design and delivery of program activities.

2. Evaluation can serve as a method *for developing policy and practice criteria* as they relate to the program sponsor's goals, by measuring program worth and cost through comparative studies of different combinations of methods for attacking specified problems.

3. Evaluation can be utilized as a procedure *for conducting research to draw universal conclusions from data* and add knowledge of principles and practice to general application in the field by determining the extent to which those conclusions from a carefully controlled situation apply to other situations (external validity).

4. Evaluation can act as a process *for intervening in an organization's procedures,* when change is the desired outcome, by involving program stakeholders in redefining responsibilities for learning through goal setting, data collection, and debriefing activities.

5. Evaluation can be a vehicle *for exerting power in a political way* within an organization, or for the group sponsoring the program, through interpretation

and use of information based on a thorough and careful study (Bramley, 1991; Burnstein, Freeman & Rossi, 1985; Phillips, 1991).

Feedback Evaluation

Private, public, and military organizations fund numerous workplace literacy programs each year. Many are pilot programs or demonstration projects that require thorough and careful examination to determine those program components and models that provide effective solutions "to increase the productivity of the workforce through the improvement of literacy skills in the workplace" (*Federal Register*, August 18, 1989, pp. 34418-34422). *Feedback evaluation* is important to sponsors and to participants involved in each workplace literacy program effort throughout the life of their pilots and projects and beyond. Evaluators both internal and external to these programs can gather and analyze data that demonstrate appropriateness of goals and outcomes, along with levels of effectiveness for program design, development, implementation, and process.

Control Evaluation

Comparing collected findings to those from other similar or dissimilar education and training programs from the same or from different organizations can result in recommendations that establish policy and practice controls for sponsors. Evaluations conducted for this purpose, which attempt to achieve a range of outcomes by manipulating variables, are referred to as *control evaluations*. Such evaluations extend beyond feedback, posing comparative questions, and tend to result in reports containing many recommendations.

Research Evaluation

Creating evaluation designs to operate in carefully controlled situations transforms the evaluation into a *research evaluation* project. Such evaluations put forth hypotheses and try to prove or disprove them, thus attempting to draw conclusions that add to the body of knowledge on workplace literacy training.

Intervention Evaluation

Using evaluation as a planned intervention to bring about change is sometimes called a "situation-specific strategy." The key steps in *intervention evaluations* are: identifying

operating systems and change networks within the organization, articulating means-ends relationships in the change process, generating profiles of information needs and evaluation issues, developing a situation-specific agenda, and implementing action-based methodologies and techniques (Bhola, 1979). This type of intervention process is similar in design to that used in changing organizations from single process, hierarchical orientations to cross-trained, flattened systems through the intervention of Total Quality Management strategies.

Political Evaluation

Using evaluation as a means for exerting power within an organization is a purpose that is almost impossible to avoid. Following the identification of indicators and their use for measuring change, the next step almost always involves setting standards. Knowledge of the degree to which individuals meet or are responsible for meeting standards can be wielded to reward, punish, or eliminate functions or workers.

The Evolution of Evaluation Methodologies

Much of the theory and techniques that exist in the modern field of evaluation have evolved from the practices used in evaluation of education in the United States. In the 1930s, strong emphasis was placed on the need for objectives by which to organize school curricula (Tyler, 1950). Objectives became the basis for planning and guiding instruction, for preparing test and assessment procedures, and for systematic program evaluation. This practice led to the comparison of programs and the development of standards by which all education programs were measured across the United States.

In the 1960s, it became apparent that the development of new curricula would need a new type of evaluation that provided feedback to program developers during the development period (Cronbach, 1963; Scriven, 1967). From this era came the terms "formative evaluation," concerned with program development, and "summative evaluation," concerned with program worth. This distinction allowed for program evaluation that focused on the performance characteristics of an individual developing program, rather than on comparison with other programs.

By the mid-1970s, a new version of independent program evaluation had developed, called "responsive evaluation," in which the evaluator is primarily concerned with program effects in relation to interests and goals of the program "stakeholders" (Stake, 1975). Responsive evaluation begins by identifying the goals each stakeholder group has for the program. This step is followed by evaluator observations of program materials and implementation, which are compared to the various goal expressions. The evaluator identifies issues and concerns, collecting data to satisfy these with a variety of instruments. Collected information is organized into themes and issues, and concerns are matched to audiences. A technique called "progressive focussing" can be used to systematically reduce the breadth of inquiry and concentrate attention on emerging issues (Parlett and Hamilton, 1977).

Evaluation of workplace literacy programs should emphasize formative evaluation while including cyclical summative evaluations, based on the premise that the workplace will continue to experience accelerating change and training in new skill applications will be an ongoing need (Philippi, 1992). Evaluation of workplace literacy programs should also measure independent program effectiveness, by comparing each program with its own stated goals, rather than comparing it with other programs. This approach is necessary because each workplace literacy program designed in response to company and worker needs is unique; no two organizations have exactly the same set of critical job tasks to be addressed by instruction (Philippi, 1993).

Responsive evaluation has potentially strong features for evaluating workplace literacy training because it attempts to take into account the interests of various groups of stakeholders, rather than just those of the program developers or sponsors, and to collect data that meet the needs of these stakeholders.

Matching Evaluation Methodology to Program Purposes

If the purpose of program evaluation is to provide decisionmakers with feedback about the extent to which a specific program is achieving the goals of all the key players, i.e., the sponsors, participants, and providers, a feedback, or *responsive evaluation,* model should be used. The goals of each group of stakeholders can be compared to actual implementation practices, use of program resources, and outcomes to determine levels of effectiveness.

If the purpose of program evaluation is to provide decisionmakers with alternatives for deciding which workplace literacy training techniques and activities are the most cost effective and successful for reaching organizational goals, a comparative, or *control evaluation* model should be used. Comparing one method of treatment to another with the same or matched populations within one organization or across several organizations could provide this kind of information.

If the purpose of program evaluation is to test a hypothesis about a particular facet of a workplace literacy training program, a *research evaluation* model should be used. This model requires careful adherence to established formal practices for structuring a situation and activities to prove or disprove the hypothesis and draw one or more universal conclusions from the results that can be added to the existing knowledge base about workplace literacy training programs.

If the purpose of program evaluation is to provide an organizational intervention to enact change, a *situation-specific strategy evaluation* model should be used. Systems inside and outside the organization that affect operations must be identified. Networks and chains of command inside and outside the organization also need to be studied to determine who must be involved and/or convinced with the data collected in order to instigate the desired change reactions. Cause-effect relationships in the change program then must be articulated before an evaluation design can be created and executed.

Responding to the Needs of Business and Industry

Because the end goal of workplace literacy training programs is to improve the skills and productivity of the workforce, programs should be viewed as an integral part of organizational training--as a long-term investment in the building of a highly skilled labor force. Most organizations evaluate their return on investment in training by using basic measurements of work output to meet product (or service) goals, emphasizing quality, quantity, variety, and/or the uniqueness of whatever is being produced or provided. Some common indicators that are measured include units produced, defects or failure rates, tasks completed, rework, scrap, backlogs, units sold, shortages, accidents, and so on.

Workplace literacy programs may also have system goals that emphasize growth, profits, modes of functioning, and return on investment. Frequently used criteria are: productivity, processing time, operating costs, efficiency, amounts of overtime or lost time, machine downtime, performance/cost ratio, et cetera. In addition, "hygiene" benefits like reduced turnover, absenteeism, strikes, and so on, plus reduction in accidents are also used to calculate return on investment or to conduct cost-benefit or cost-effectiveness analyses (Cascio, 1982; Kearsley, 1982). When data are already being collected for such measures, employers and employees may wish to examine them to identify relationships exhibiting positive changes in performance.

Measures of individual productivity are easiest to correlate statistically, but it is difficult to isolate training impact from other variables that may be producing a cost benefit. Additionally, benefits can be rather diffuse and often take considerable time to be realized, which makes cost analysis somewhat inappropriate as a tool for evaluating training and development activities (Bramley, 1991).

An easier means for costing changes in the workplace directly resulting from a training program can be found in the process of evaluating the *value added to employees.* The value-added approach assumes that there is a need for certain skills in areas relevant to the work performed, that there is motivation to do the job, and that there is an opportunity for the skills to be used to perform the job.

Pretraining analysis must be conducted to estimate the position of the individual targeted participants with respect to various levels of skill and motivation of typical employees who are doing that kind of work. If employees are thought to have the motivation but lack some important skills, then the investment should be in employee skills training. Postprogram measures of skills and motivation are also collected and changes calculated. Assuming that managers and supervisors can accurately rate their employees, that there is opportunity and encouragement to use the new skills, and that the skills required for successful job performance have been accurately identified and are being taught in training, value-added evaluation can be used to estimate the return on training investment (Philippi, 1993).

Unlike cost-benefit analysis which summarizes all outcomes in monetary terms and omits those things that cannot be expressed in dollars or other currency units, the value-added

process describes outcomes in their own terms and then concludes that these do or do not imply significant increase in skills or motivation to perform work (Bramley, 1991).

Evaluation design, as a part of workplace literacy training programs, needs to determine program effectiveness according to the value yardsticks that business and industry apply to training. Quality indicators, measures, and performance standards for individual programs should be identified from the workplace community, rather than from educational environments.

Whenever companies evaluate training, they generally assess effectiveness on four different levels:

1. To what extent does the proposed training program match a critical organizational training need identified by the employer or employees?

2. To what extent do employees who participate in the training master the content of the training program?

3. To what extent does mastery of training transfer to and positively impact on job performance?

4. To what extent does changed job performance result in cost benefit to the organization?

These evaluation levels differ significantly from traditional education evaluation criteria. Much of evaluation of educational programs focuses on individual progress toward universally defined standards, i.e., grade levels or letter grades, which sometimes are then compared to group achievement. One factor that is shared between organizational evaluation of training and educational evaluation of school-based programs is the eliciting of feedback from individual participants and instructors to determine the need for adjustment to materials, scheduling, or instructional techniques. In business and industry evaluations, such feedback is generally solicited as a structured, formal activity built into each training program.

Expressed in other terms, the four levels of training evaluation are often referred to by human resource development personnel as the Kirkpatrick Approach (Kirkpatrick, 1983). Based on a conceptual framework for classifying areas of evaluation, the model lists the four levels as:

Reaction-- Were the participants pleased with the program?

Learning-- What did the participants learn in the program?

Behavior-- Did the participants change their behavior based on what was learned?

Results-- Did the change in behavior positively affect the organization?

Several adaptations of this classification have been used and published by Xerox, AT&T, IBM, and Saratoga Institute (Phillips, 1991). In any of these variations of the categorical levels of organizational evaluation for training, all four levels are always addressed in the evaluation of every training program.

A Model for Evaluating Workplace Literacy Training Programs

A model that is reported by the American Society of Training and Development (ASTD) to be receiving widespread use in a number of organizations is the *Context, Input, Process, and Product* (CIPP) evaluation model (Galvin, 1983). Developed in 1971 by educational researchers Daniel Stufflebeam and Egon Guba as part of Phi Delta Kappa's National Study Committee on Evaluation, the model is designed to provide feedback to program decisionmakers (Bhola, 1979; Galvin, 1983). It meshes well with the four levels of organizational training evaluation and with the value-added training evaluation model.

Originally put forth as a means for evaluating reading programs, the CIPP model provides a framework for clarifying program goals and objectives, observing whether or not they have been achieved, and producing information that can be used by program decisionmakers to improve the program's capacity to achieve its goals. This model has been used to evaluate numerous workplace literacy training programs across the country and could function as a uniform method for comparative and research evaluations along with its current role in individual program feedback evaluations.

The model poses a series of research questions in each of the following four areas:

Context-- Attempts to review and clarify the underlying philosophy and goals of a workplace literacy training program. Answers the question, *To what extent are the goals of key program players congruent or divergent?*

Input-- Determines whether or not required resources exist. Answers the question, *To what extent are program resources adequate and appropriately utilized?*

Process-- Compares program activities to program goals. Answers the question, *To what extent are program development and implementation being carried out according to program goals?*

Product-- Attempts to examine program outcomes as defined by program objectives for participants and for the organization. Answers the question, *To what extent are program goals being met?*

Recommendations for Designing Workplace Literacy Evaluations

Planning effective procedures for conducting evaluation activities requires careful and thoughtful construction of an evaluation design. The process for evaluation design involves several distinct steps. After the purpose for the evaluation has been determined and a model selected, issues should be identified, relationships defined, and investigatory techniques selected.

The Design Process

Thinking through all of the possibilities for formulating an evaluation design is an important initial phase in developing the plan. Deciding which issues should be investigated and what strategies can be employed to acquire the necessary information from which conclusions can be drawn is usually the first step. These decisions often must then be modified and reworked because of situational or programmatic limitations.

Data collection methods should be designed to elicit the maximum amount of appropriate information about the program with a minimum of interference in everyday worksite operations. A time line or schedule of evaluation activities must be constructed to ensure the timeliness and thoroughness of the investigation and reporting of information.

Identifying Issues for Investigation

Each workplace literacy training program evaluation has specific individual program issues that it must respond to. Questions should be identified that can be answered by the collection and analysis of the information gathered during the investigation.

For most programs the main issue facing program decisionmakers is w ether the level of program effectiveness warrants keeping the pilot program in the budget as a continuation or expansion. This decision can be especially critical if the pilot phase of the program has been operated with support monies from an external funding agency (e.g., from a federal government grant or other limited funding source).

Questions asked by an evaluation should consider more than reactions of participants and trainers or levels of mastery of training content. Program decisionmakers need hard data to determine the extent to which the program has had a positive impact on performance of job tasks that have been identified by the organization and its workers as critical. Selecting local performance indicators can assist evaluators with the formulation of specific questions to be investigated by identifying measurable behaviors and desired levels of positive change to be anticipated over a set time period.

Determining Relationships

The process for designing an evaluation plan also should examine what relationships exist among the categories that have been selected for inclusion. For example, if the CIPP model were used, the evaluation would need to consider the relationships between goals and resources, resources and development and implementation processes, and among resources, implementation, and outcomes. This approach aids with recognition of multiple variables that may have a direct or indirect relationship with each other and an impact on findings.

Sticht (1991) points out that there is often a basic assumption made that there is a relationship between various literacy abilities and job productivity. It is important to note that this is not true for all aspects of productivity. He explains that many job tasks do not require the direct application of reading, writing, math, or communication abilities and can be learned by watching others. For this reason, it is critical that an evaluation take into account the role of literacy ability in relation to various indicators of productivity. Otherwise, it may not be possible to demonstrate that the program has, in fact, increased productivity. Determining relationships of program components to each other serves as a procedure for focusing the investigation.

Establishing Indicators, Measures, and Standards

Once investigation topics have been identified, *quality indicators* should be established within each area. A quality indicator is a variable that reflects effective and efficient performance (Condelli, 1992). For example, if an applied math skills program were taught to support the use of Statistical Process Control (SPC) in the workplace, a quality indicator might be a worker's ability to accurately plot production monitoring data on a run chart.

A *performance measure* is the data used to determine the quantitative levels of performance. It is the operational definition of the indicator (Condelli, 1992). For example, the measure for the indicator of accurately plotting production monitoring data on an SPC run chart might be the number of errors found on data charts plotted by a worker.

A *performance standard* is a measure with a specific numeric criterion or level of performance tied to it. Standards define levels of acceptable performance on the measure. They may be established for a single point in time or to measure increases in performance over time (Condelli, 1992). For example, a standard may be set for the maximum number of errors acceptable in plotted data on SPC run charts; or a measure may be established as a preprogram standard to be compared to a postprogram measure to determine if a new standard for fewer errors can be set.

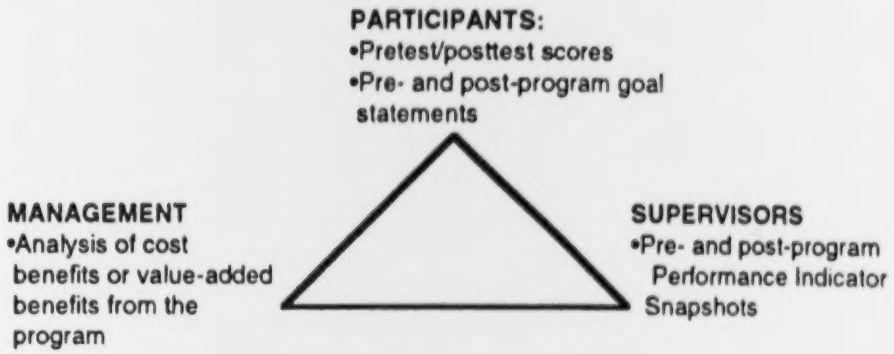
Quality indicators provide precise vehicles for evaluating the success of a program in meeting its stated goals. Performance standards are more difficult to use because they compare program results to preconceived or preestablished levels of performance. For individual programs concerned with feedback on achievement of goals, they are primarily of use in comparing one program cycle to others over time.

Planning for Input from Multiple Sources

The way data are collected is also critical in any evaluation. The methods and instruments used, and the thoroughness and completeness of information gathered, can determine whether or not findings are valid and convincing. Data should be collected from many different sources to corroborate the reliability of information obtained and to draw comprehensive conclusions about the evidence. A variety of techniques can be employed to gather information: structured interviews, surveys, focus groups, pretests and posttests, and observations. If instruments or forms need to be created, it is important that the format chosen allows them to be simple to complete and easy to analyze.

A research method called *triangulation* can be used to validate evaluation findings. Simply stated, "triangulation" is the collection and comparison of information from three or more diverse sources to determine similarities or differences in evidence of program effectiveness (Philippi, 1991). To compare and examine the findings from different points of

view in a workplace literacy training program, data should be collected from the sources specified below:



Additional information about program effectiveness also should be collected from key program staff, instructors, curriculum developers, and advisory committee members (see Figure 1).

Triangulation of data should be attempted with each category of evaluation. For example, if the CIPP model is used to evaluate a workplace literacy training program, information from multiple diverse sources should be gathered and compared for Context, or goal statements; for Input, or resources; for Process, or development and implementation; and for Product, or outcomes. In order to achieve triangulation of Context data, statements of program goals could be obtained from management, instructors, program developers, curriculum writers, participants, supervisors of participants, and any written documents containing published program goals. For Input data, information about adequacy and appropriate use of program resources could be gathered from instructors and participants in regard to materials and facilities, from instructors and writers in regard to time and support, and from managers and supervisors in regard to costs and release time, if applicable. Data for Process and Product categories could be triangulated in similar ways.

A Menu of Collection Techniques

A variety of data collection techniques should be selected for use with any workplace literacy training program evaluation. Methods should be chosen that are most advantageous

Figure 1

The Who, What, When, and Why of Evaluation

Evaluation Levels	Evaluation Strategies
Level 1. To what extent does the proposed workplace literacy training program match a critical organizational training need?	• Work with employers and employees to determine organizationally critical job tasks and program goals. • Identify the skill applications needed to perform those tasks competently. • Collect pre- and post-program data on employee expectations and satisfaction with program content and delivery.
Level 2. To what extent do employees who participate in the training master the content of the program?	• Collect representative examples of participants' work to demonstrate progress toward program instructional goals. • Develop and administer competency-based, parallel pretests and posttests correlated with program goals and content to determine progress of participants toward mastery of instructional content.
Level 3. To what extent does mastery of training transfer and impact positively on job performance?	• Identify local indicators for measuring performance of specific behaviors on critical job tasks. • Collect ratings of performance, either by participants' supervisors or as employee self-ratings, on identified indicators pre-and post-program course cycles to measure the amount of transfer from training to job performance.
Level 4. To what extent does changed job performance result in cost benefit to the organization?	• Guide employers and employees to examine areas in which critical job tasks are performed to determine the extent to which changes in performance after participation in workplace literacy training programs are saving or generating money for the organization. • Measure skill and motivation levels in performance of job tasks by targeted program participants before and after training to conduct a study of the value added to employees by the training.

Note: Adapted from Philippi (1992).

BEST COPY AVAILABLE

to a program's particular circumstances and data needs. Some suggested methods for collecting data are described below.

Surveys and questionnaires. These can be highly effective to use when responses are required from large numbers of people. They have the advantage of retaining consistency during questioning of all respondents. The greatest disadvantages with their use are poor rates of return without persistent follow-up and difficulty with developing questions that do not "beg the answer," confuse, or bias respondents. Surveys might be used in preprogram information gathering, as well as for postprogram ratings. The main purpose of a survey used with program participants should be to obtain information about learner goals and expectations; this information can be compared to postprogram learner reactions to determine extent to which goals were perceived to have been met. This comparison can also be a good way to determine the appropriateness of course publicity, content, and recruitment procedures. Asking open-ended questions can provide quotations to illustrate conclusions in an evaluation report. In addition to questions relating directly to the workplace literacy training program, participant surveys can be useful vehicles for gathering demographic data about participants' age, gender, ethnic group, education level, job experience, and so on. This information should enable the development of a composite learner profile for participants in the program. By comparing such a profile to similar statistics on an average worker profile for the organization or for the worksite, the suitability of the program materials for use with other workers or departments can be determined.

Polarized scales. If surveys and questionnaires become too demanding or lengthy, they are often returned with less than complete information. If time is an issue for respondents who are filling out the instruments, the use of polarized scales to collect the information might be considered. On polarized scales, respondents are asked to mark numbers to indicate the strength or weakness of their feelings or preferences. This technique can be a quick way to collect data about the effectiveness of instructional materials, or ratings of participant performance by themselves or by their supervisors. Numerical scales also can provide a means for quantifying data. Percentages can be calculated for the number of respondents who circled each of the number choices. The numbers can be arranged as a continuum or as denoters of categories.

Structured interviews. A series of questions can be formulated to be asked of respondents orally. Secondary probing questions can be created for each main inquiry, in case more detailed information needs to be elicited from respondents about a topic. All of the questions should be written out before the interview and read to respondents one-on-one or in small groups. Nothing more should be said to expand on each question read by the evaluator. Respondents are given time to answer and their responses are recorded. Each time a structured interview is conducted, the questions it contains should be read in the same manner with no deviations from what is scripted, so that all respondents can be given consistent information to react to. Structured interviews may also be conducted with slightly larger focus groups.

Observations. This technique can be especially appropriate for use in collecting program implementation data. Recording of all trainer and participant actions and behaviors observed in a session as they occur, plus the time each instructional activity began and ended, yields an objective account of the instructional process. Gathering this type of information from several training sessions across time, or from the same session conducted by different trainers, can provide an avenue for the evaluation to generalize about the amount of time spent in various types of instructional activities and whether instructional materials are being presented as designed. Observations can provide insights into the effectiveness of instructor training, as well. They offer numerous ideas for ongoing inservice and staff support throughout the life of a program. The observer's recorded notes from the instructional session also can provide an excellent feedback vehicle for discussing performance or other delivery issues with an instructor immediately following the observation.

Gathering Baseline Data

Before implementing the delivery of workplace literacy training, it is critical to collect data pertaining to the existing conditions. Such information is necessary to determine whether change has occurred. For example, if a program posttest is going to be administered to demonstrate level 2 mastery of training content, a pretest should also be administered before actual skill instruction begins. The results of the pretest document, or *benchmark*, existing conditions before treatment. For example, if a participant scores high on a posttest, it does not necessarily mean that the score is the result of what he or she learned during

training; the participant may have entered training with the ability to achieve a high score on the posttest. There is no way of knowing which situation the posttest is evidencing without administering a pretest; only by comparing pre- and posttest scores can evidence be obtained that may be used to demonstrate gains due to training.

It is also important to collect baseline data about goals before and after program cycles. Goals can change in relation to conditions external to the program and should be monitored throughout the program life. Baseline data collection is important in terms of measuring level 3, impact on performance, too. Preprogram levels of performance should be established in order to demonstrate that any change has or has not occurred for the indicators selected. Gathering data to document the conditions that exist prior to program implementation provides a point of reference with which to compare all postprogram results.

Reporting Findings

Evaluation reports should be organized around the information needs of the users. Interim reports should be concise and formatted to facilitate quick access to observations, conclusions, and suggested next steps. Summative reports should be issued at the end of training cycles. These reports should be much more detailed and contain an objective description of data collected, conclusions reached from data analysis and interpretation, and recommendations based on findings. In compiling information for an interim report, addressing current issues or providing progress updates is usually the main purpose for generating the reports. Of secondary importance, although sometimes necessary, can be the use of interim reports to create documentation of the evaluation process. The program stakeholders should be the recipients of interim reports. Their concerns should be addressed by the contents. Final reports produced at the end of program cycles may have a wider and more diverse audience. Because some of the readers of a summative report may be less familiar with program and evaluation details than recipients of the interim reports that preceded it, the final report should contain a description of the program and evaluation design along with data, conclusions, and recommendations.

Summary

Each workplace literacy training program designed in response to company or worker needs is unique because no two organizations have exactly the same set of critical job tasks to be addressed by instruction. For this reason, evaluation of workplace literacy training programs needs to measure program effectiveness independently, that is, by comparing each actual program to its own stated goals, rather than comparing it to other training programs. At the same time, evaluations need to be conducted in such a way that it is possible to draw universal conclusions from data that may be added to the base knowledge for the field.

Evaluation may be conducted for a variety of reasons that include provision of feedback to program decisionmakers for program quality control; comparisons of techniques and variables to control program criteria; research to prove or disprove specific hypotheses; intervention in an organization to promote change; and exertion of political power within an organization. Identifying the purposes for evaluation and selecting a model and methodology to achieve those purposes are critical to obtaining the desired information.

Of particular importance in designing an evaluation model is consideration of the questions for which the evaluation will contribute information that helps determine answers, the relationships of program activities and goals to those questions, and techniques that can obtain enough evidence to draw conclusions about the program areas under investigation.

Selecting local program indicators and developing data collection instruments and procedures for measuring changes in those indicators are the core of any good workplace literacy training program evaluation. Indicators should be workplace specific. Analysis, interpretation, and reporting of data obtained should be in ways that are meaningful to program decisionmakers at all levels.

References

Bhola, H. S. (1979). *Evaluating Functional Literacy*. Tehran, Iran: Hulton Educational Publications.

Bramley, P. (1991). *Evaluating Training Effectiveness: Translating Theory into Practice*. London: McGraw Hill.

Burnstein, L., Freeman, H. E., & Rossi, P. H. (1985). *Collecting Evaluation Data*. Beverly Hills, CA: Sage.

Cascio, W. (1982). *Costing Human Resources: The Financial Impact of Behavior in Organizations*. Boston, MA: Kent.

Condelli, L. (1992). *Quality Indicators, Measures and Performance Standards*. Washington, DC: U.S. Department of Education, Office of Vocational and Adult Education.

Cronbach, L. (1963). Course Improvement Through Evaluation. *Teachers College Record, 64*, 672-683.

Fields, E. (1976). Industry Based Programs: A Growing Source for Adult Literacy Development. *Lifelong Learning, 10*, 7-9.

Galvin, J. G. (1983). What Can Trainers Learn from Educators about Evaluating Management Training? *Training & Development Journal*, 52.

Kearsley, E. (1982). *Costs, Benefits and Productivity in Training Systems*. Reading, MA: Addison-Wesley.

Kirkpatrick, D. (1983). Four Steps to Measuring Training Effectiveness. *Personnel Administrator*, November, 1983, pp. 19-25.

Parlett, M., & Hamilton, D. (1977). Evaluation as a New Approach to the Study of Innovative Programmes. In D. Hamilton, D. Jenkins, C. King, B. MacDonald, and M. Parlett (Eds.), *Beyond the Numbers Game*. London: Macmillan.

Philippi, J. (1991). *Literacy at Work: The Workbook for Program Developers*. New York: Simon & Schuster.

Philippi, J. (1992). *How Do You Know If It's Working? Evaluating the Effectiveness of Workplace Literacy Programs*. Washington, DC: U.S. Department of Education, Office of Vocational and Adult Education.

Philippi, J. (1993). *Retraining the Workforce: Meeting the Global Challenge - Module VI: Measuring Program Effectiveness*. Dallas, TX: Dallas County Community College District.

Phillips, J. (1991). *Handbook of Training and Measurement Methods* (2nd ed). Houston, TX: Gulf Publishing.

Scriven, M. (1967). *The Methodology of Evaluation* (AERA Monograph Series in Curriculum Evaluation). Chicago, IL: Rand McNally.

Stake, R. (1975). *Evaluating the Arts in Education: A Responsive Approach*. Columbus, OH: Merrill.

Sticht, T. (1991). *Evaluating National Workplace Literacy Programs*. Washington, DC: U.S. Department of Education, Office of Vocational and Adult Education.

Tyler, R.W. (1950). *Basic Principles of Curriculum and Instructional Design*. Chicago, IL: University of Chicago Press.

ADVANTAGES AND DISADVANTAGES OF LONGITUDINAL DESIGNS FOR EVALUATING WORKPLACE LITERACY PROGRAMS

Elisabeth Hayes
University of Wisconsin-Madison

Abstract

A basic characteristic of longitudinal designs is the collection of data over time. Advantages of longitudinal designs for program evaluation include the potential to assess: (a) outcomes that develop only over time; (b) the stability of program outcomes over time, and (c) patterns of program impact over time. Longitudinal designs have general methodological disadvantages as well as disadvantages for workplace programs in particular. Longitudinal studies that incorporate elements of experimental designs offer the greatest potential for establishing relationships between program participation and observed changes. However, experimental procedures can be difficult to implement in the context of workplace programs.

Introduction

Longitudinal impact studies of adult literacy education are increasingly advocated as part of national efforts to evaluate program effectiveness (see, for example, Pelavin Associates, Inc., 1992). While adult literacy education may help students achieve many immediate goals, some outcomes, such as employment or further education, may only be realized over longer periods of time. As Beder (1991) observes, "while the short-term effects of adult literacy education may be modest, the long-term effects compounded over a lifetime may be enormous" (p. 156). Workplace literacy programs in particular may have an impact on workers' productivity and job advancement that can be fully measured only over time.

The assessment of such long-range outcomes may provide important justification and ongoing support for workplace programs. However, there are a number of problems associated with longitudinal studies that may limit their utility for workplace program evaluation. The purpose of this paper is to discuss both advantages and disadvantages of longitudinal designs in relation to the evaluation of workplace literacy programs. Several alternative longitudinal designs are described to illustrate strengths and weaknesses of different methodological approaches.

What is longitudinal research?

Longitudinal designs have certain characteristics that distinguish them from other research designs. Simply put, longitudinal designs involve the collection and analysis of data over time (Lewis-Beck, in Menard, 1991). A more complex definition offered by Menard (1991) includes three elements:

> Longitudinal research is research in which (a) data are collected for each item or variable for two or more distinct time periods; (b) the subjects or cases analyzed are the same or at least comparable from one period to the next; and (c) the analysis involves some comparison of data between or among periods. (p. 4)

Longitudinal designs may be contrasted with simple **cross-sectional research** in which data are collected at one point in time. They also differ from **experimental designs** in that longitudinal studies do not necessarily involve the use of control or comparison groups and

random assignment. However, experimental methods can be used with longitudinal designs to better answer questions about the long-term effects of educational interventions.

Advantages of Longitudinal Research

Several key advantages of longitudinal designs are based on their potential for answering questions about program outcomes that are not answerable by other (possibly more feasible or less costly) designs. Specifically, longitudinal designs can assess (a) program outcomes that develop only over time; (b) the stability of program outcomes over time, and (c) patterns of program impact over time.

For the purpose of this discussion, I will use three general types of learner outcomes as examples of outcomes that might be important in evaluation of workplace literacy programs. These outcomes are adapted from Mikulecky (1992, p. IV-3). The first type is **literacy-related beliefs and abilities**. Literacy-related beliefs include perceptions of literacy, oneself as a literate individual, and aspirations toward becoming more literate (Mikulecky, 1992). Abilities include general reading, writing, and mathematical abilities as well as more specific task-related competencies. Examples of the second type, **productivity outcomes**, as described by Mikulecky (1992) include improved job performance, attendance, safety, productivity suggestions made and approved, and disciplinary measures and grievances. The final type of outcome, **transfer/carry-over**, includes job-related outcomes such as job retention, job advancement, and higher income. Other outcomes may be an increase in family literacy activities or plans for and actual participation in further education and training.

Program Outcomes That Develop Over Time

Longitudinal designs permit the identification of outcomes that may not be evident or measurable for all participants at one point in time (for example, when data are typically collected at program completion). Longitudinal designs ar· perhaps least essential for assessment of changes in literacy-related abilities and beliefs. Typically, program impact on literacy skills and beliefs can be assessed at the end of participation. However, longitudinal designs may be more important to assess the use of newly acquired skills on the job, which

might become most apparent over time as various opportunities arise for the application of skills.

In contrast, longitudinal designs appear to be quite important to assess changes in productivity and transfer/carry-over. It may take an extended period of data collection to determine whether a worker is more productive or has better attendance. Job mobility or participation in further education might also occur only after an extended period of time.

Stability of Program Outcomes

Longitudinal designs can also provide useful information about the stability of program outcomes. While programs tend to be based on the assumption that literacy skills are maintained once acquired, there is some evidence that general literacy gains are quickly lost unless they are practiced and reinforced on the job (Mikulecky, 1992). The stability of productivity gains might also be most effectively determined through longitudinal studies. It is possible, for example, that an individual's productivity might increase temporarily during or at the end of program participation, perhaps as a type of Hawthorne effect, but then decrease as time passed. The same phenomenon might be true for transfer of skills, such as in family literacy activities.

Patterns of Program Impact Over Time

In addition to providing information about simple stability of outcomes, longitudinal studies can provide insight into patterns of program impact over time. For example, literacy skills and perceptions of oneself as literate may continue to develop over time as the learner applies those skills in daily situations. Patterns in improved job performance, possibly related to fluctuations in on-the-job demands, may be necessary to track over time, to determine overall program impact in the workplace. Longitudinal designs may be particularly important to assess patterns of program impact on other outcomes, such as long-term income gains, over time. There might be an apparently small gain over one or t o years, but after a number of years the gains might multiply rather than be simply additive

In general, the advantage of longitudinal designs is the potential to obtain a more comprehensive and extensive picture of the outcomes that might be related to participation in

workplace literacy programs than can be gained through a single follow-up study or an evaluation that ends when the learner leaves the program.

Disadvantages of Longitudinal Designs

This section will identify and briefly discuss two general types of disadvantages: (a) methodological disadvantages of longitudinal studies in general, and (b) disadvantages of longitudinal studies for workplace programs in particular.

Methodological Disadvantages

Five methodological disadvantages of longitudinal designs seem to be particularly salient to the workplace literacy context. These include respondent attrition, respondent conditioning, unreliability of measures, changes in measurement over time, and impact of other individual and social factors.

Respondent attrition. One of the most commonly mentioned problems in longitudinal studies is the loss of respondents over successive periods of data collection. In workplace studies, attrition might result from respondents' leaving their job with no forwarding address or simple refusal to participate if follow-ups are perceived as increasingly intrusive, time consuming, or threatening.

Respondent attrition has at least three negative effects on a program evaluation. First, it suggests the possibility of **systematic bias** in those who drop out compared to those who remain in the study (Menard, 1991). For example, respondents who drop out may be those who are successful in getting better jobs, so evidence of positive impact is diminished. On the other hand, individuals who are less successful in their jobs may be more likely to move, or individuals may leave for reasons unrelated to possible program outcomes. Certain kinds of bias can be identified through collection of information on why respondents drop out of the study, but it may not be possible to assess all relevant differences in respondents and dropouts. Loss of respondents also can result in **nonequivalent groups**, or groups that are not really comparable over time when control or comparison groups are used in the study design (Mark & Cook, 1984).

In addition, an appreciable decrease in sample size can lead to **reduced analytic power**. The smaller the sample, the more difficult it is to establish significant relationships among variables or differences between groups. The likelihood increases that findings will be affected by individual variability or error in measurement. This may be a particular problem for workplace literacy programs since participant samples tend to be quite small. Mikulecky's N.C.A.L. study, for example, used groups of only 12-15 students (Mikulecky, 1992); with such small samples, the loss of only a few respondents greatly reduces the potential to identify significant changes.

These problems suggest that strategies such as the use of incentives and other means of maintaining participation may be critical to the success of a longitudinal design.

Respondent conditioning. A possible effect of repeated measurement of the same variables in longitudinal studies is that respondents may learn to anticipate questions and answer them in a certain way (Menard, 1991). They may simply repeat previous responses without thinking, or be reluctant to describe changes that might be perceived as undesirable, such as poorer job performance. In addition, respondents' behavior and attitudes may actually change as result of being involved in the study rather than due to their participation in the educational program alone. For example, questions about use of work-related literacy skills may increase individuals' awareness of their literacy practices and improve their performance on the job. It can be difficult to separate effects of participation in the study from program outcomes.

Unreliability of measures. Another problem in longitudinal research is the possibility that observed changes might be due to errors in measurement rather than real change. For example, many measures of self-concept have been found to be unreliable; a person's responses can be affected by mood swings or the conditions in which the questions are administered, leading to incorrect indications of change. Unreliability is also a problem with commonly used tests of literacy ability. For example, grade level results of the Test of Adult Basic Education (TABE) can be affected by correct or incorrect responses to one or two questions.

Changes in measurement over time. Data collected at different times in longitudinal designs can be affected by unintentional changes in who is involved in assessing outcomes and how outcomes are assessed. For example, one possible measure in a workplace literacy

program evaluation is supervisor assessment of worker productivity. Supervisor assessment might be affected by factors such as the supervisor's increased awareness of worker performance or a change in the supervisor's standards over time. If supervisors change, the new supervisor may have a different way of assessing productivity, or have no basis of comparison for assessing change. To minimize such problems, data collection must be closely monitored over time, and in particular consistent training of those responsible for data collection is necessary (Menard, 1991).

Increasing impact of other individual and social factors. Over time, it becomes increasingly difficult to separate program effects from the effects of other events and experiences. In the workplace, for example, people may receive pay increases over time regardless of whether their performance has improved. The nature of the economy is a particularly important consideration in assessment of workplace literacy program impact. A period of high unemployment might make it likely that individuals will be unable to obtain new jobs and perhaps be more motivated to pursue further education. Such factors may be partially accounted for through the use of control or comparison groups, although this method obviously increases the complexity of the evaluation design and analyses.

In general, these issues reflect one key disadvantage of longitudinal studies for program evaluation: the increasing difficulty of establishing cause and effect relationships over time. Experimental designs and other strategies can be used to minimize some of the above problems, but such strategies also can become increasingly difficult and costly to maintain over time.

Specific Disadvantages for Workplace Literacy Programs

Longitudinal evaluation designs also present particular disadvantages for workplace literacy programs. These disadvantages include loss of learner anonymity, loss of voluntary education participation, increased administrative demands, maintenance of support, and inappropriate outcome expectations.

Loss of learner anonymity. Many workplace programs protect the identity of participants as a condition of program operation (Mikulecky, 1992). Anonymity may be considered important to protect employees' job security and minimize the disclosure of skill deficiencies to co-workers and supervisors. However, the need to track respondents over time

and obtain documentation of job performance increases the likelihood that anonymity will be lost.

Loss of voluntary educational participation. This problem is primarily an issue for studies that involve random assignment of individuals to participant or control groups (typically necessary to minimize differences between groups). This procedure can create "artificial" program conditions for evaluation. If some individuals are involuntary participants, they may have reduced motivation to learn, higher attrition rates, and lower achievement rates than the typical voluntary participant, leading to an inaccurate portrayal of actual program effectiveness. The possible denial of service to potential learners may be resisted by both learners and program personnel.

Increased administrative demands. Adult literacy programs in general tend to have poor record-keeping systems (Young, 1992). In addition, many employers do not maintain worksite productivity records. Accordingly, it may be necessary for programs to initiate such data collection procedures as part of the evaluation study. In open-entry/open-exit programs, respondent selection is most appropriately ongoing (a time-specific sample is likely to be biased), so data collection must also be ongoing. This increases the intrusiveness of the study and the likelihood that staff will need to be continuously involved.

Administrative demands may also be high as a result of efforts to obtain adequate samples. It may be necessary for programs to collect data for new participants over extended periods of time. As an example, the study described by Hargroves (1989) required a 10-year period to obtain about 300 participants! Programs may need to implement and maintain records for long periods of time. In addition, the high staff turnover typical of adult literacy programs (Young, 1992) may require ongoing staff training to ensure appropriate record-keeping.

Maintaining support. Another challenge for program staff and evaluators is obtaining and maintaining support for a long-term evaluation project. The cooperation of employers, supervisors, and unions as well as program participants is necessary for ongoing data collection. Maintaining the interest of these varied groups and ensuring their commitment to the evaluation over a long period of time may be problematic.

Inappropriate outcome expectations. As observed by Mikulecky (1992), individual learning time in most workplace programs is very limited, typically less than 100 hours per

year. Instruction in a number of programs focuses on specific job-related tasks rather than broader skill development. This type and amount of instruction are generally insufficient for improvements in general literacy abilities and possibly for other lasting impacts. Logically, it may be inappropriate to expect that long-term outcomes will be a result of such programs. Longitudinal studies may simply be inappropriate in relation to the goals and purposes of some programs.

Alternative Designs

In this section, I briefly describe three alternative longitudinal designs. The designs are represented graphically in Figures 1, 2 and 3. The primary difference among these alternatives is the extent that they also incorporate elements of experimental designs. In general, as more rigorous experimental procedures are used, it becomes more possible to establish relationships between program participation and observed changes. However, experimental procedures also will be more difficult to implement.

Each of the following designs starts with a representative sample of programs; procedures for selection of such a sample are not discussed here.

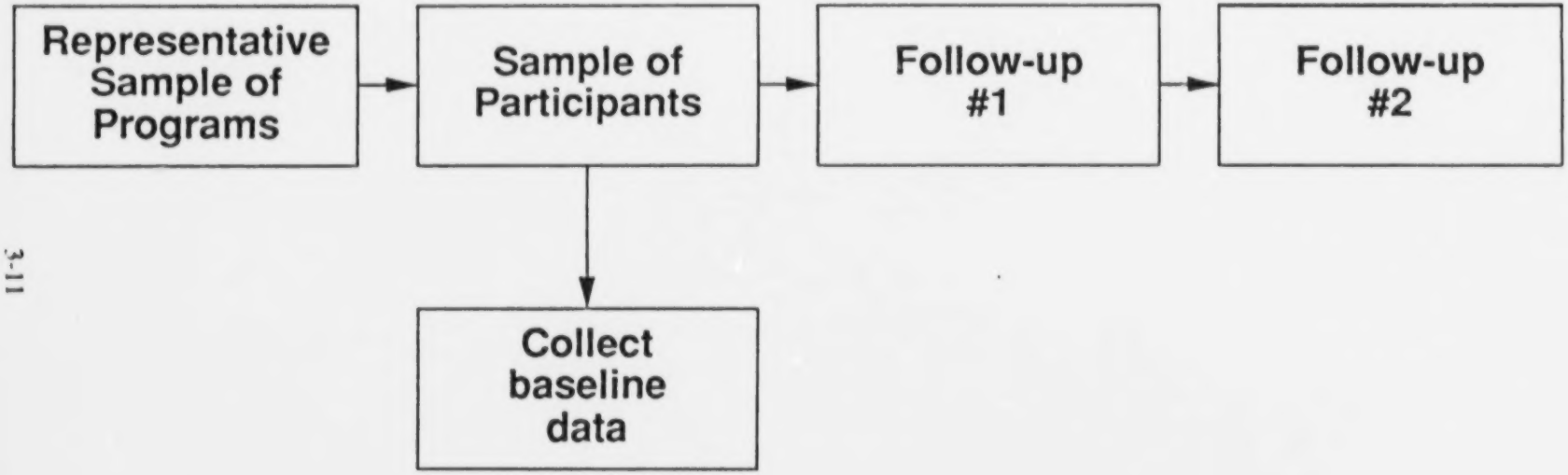
Design 1 (Figure 1)

This design includes collecting baseline data from a sample of participants and follow-up data collection at two or more later time periods. The sample might be selected randomly or through stratification according to pertinent characteristics. The intervals between follow-ups would depend on the outcomes of interest in the evaluation. This design maintains voluntary program participation and offers relative ease in sample identification. The design allows comparisons among programs or participant groups to determine relative impact of different programs or for different groups. However, no comparisons can be made between participants and nonparticipants; thus it may be difficult or impossible to establish actual program impact related to certain variables.

A variation on this design involves multiple data collection points before as well as after training. In this approach, a **simple interrupted time series design** (Mark & Cook, 1984), data collected prior to training can be used to identify any patterns of change that

Figure 1

LONGITUDINAL EVALUATION DESIGN #1



3-11

might occur without training. In effect, the treatment group serves as its own control group. This approach is not without its limitations (see Mark & Cook, 1984), but it has a number of advantages in comparison to a single assessment of baseline data. An obvious challenge is identifying and collecting data on potential trainees for a sufficient length of time prior to their participation in training.

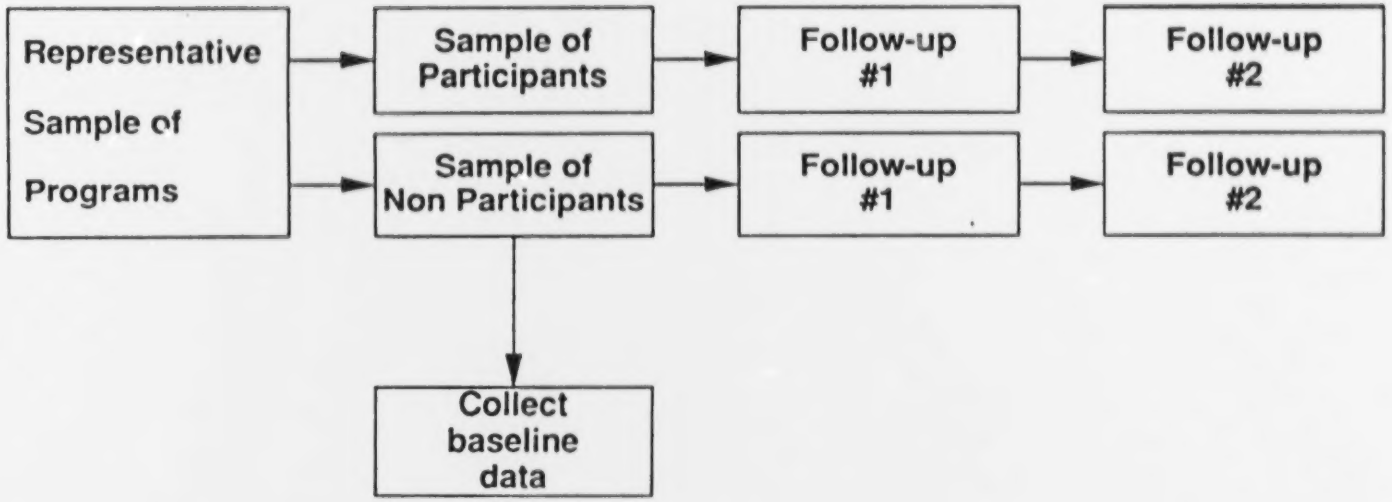
Design 2 (Figure 2)

This design includes two groups, a sample of participants and a sample of nonparticipants. The nonparticipants may be selected randomly, or matched to participants to ensure comparability on relevant characteristics such as educational level or work experience. Baseline and follow-up data are collected for both groups at the same intervals. This design is similar to the approach used in a study done by the Federal Reserve Bank of Boston (Hargroves, 1989). In this study, participants in the bank's skills development program were compared on certain indicators of job success to another group of entry-level employees. The comparison group was not equivalent; in fact they had higher levels of education and job experience that had made them ineligible for the skills program. However, the study still yielded useful information. The findings indicated that participants, who ordinarily would be expected to demonstrate poorer job performance, actually performed as well as or better than the comparison group.

This design also maintains voluntary participation in the educational program and offers relative ease in identifying samples. The comparison group offers the additional advantage of serving as a control for situational factors such as economic conditions or changes in the workplace. However, the participant and nonparticipant groups may differ in certain ways that can contribute to incorrect conclusions about program impact. For example, individuals who voluntarily participate in an educational program may be generally more motivated to achieve than nonparticipating co-workers. In the Boston bank study, a number of individuals who entered the skills program dropped out prior to completion, and weren't included in the study. Thus, the participant group consisted of more motivated, and probably more skilled, individuals than the original group--and perhaps those who would have been more likely to succeed without any educational support. This design does not make it

Figure 2

LONGITUDINAL EVALUATION DESIGN #2



possible to draw conclusions about what participants might have accomplished without any involvement in the program.

Design 3 (Figure 3)

This design, while the most complicated, has the greatest potential for providing valid information about program impact. From the representative sample of programs, the entire population of eligible persons is identified. Eligibility might be determined through basic skills testing, level of educational attainment or whatever means is appropriate for the programs. Next, eligible individuals are randomly assigned to instruction or to a control group that does not receive instruction. Baseline and follow-up data are again collected at comparable intervals for both groups.

This design offers the greatest control for possible effects of extraneous factors, and thus the strongest evidence of program impact. It removes the possible bias related to voluntary program participation. Because individuals are randomly assigned to instruction or control groups, presumably the groups are equivalent. However, the loss of voluntary participation, as discussed above, if it does not reflect normal program conditions, may yield inaccurate information about real program impact. In addition, attrition from either group will reduce their comparability and potentially invalidate the design.

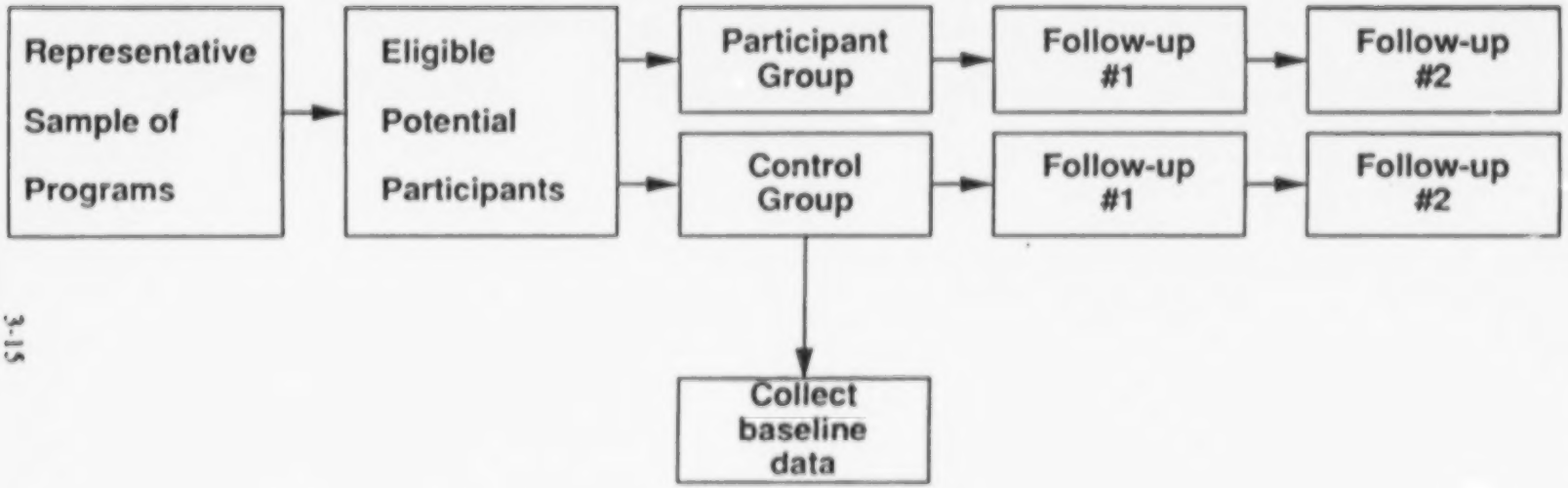
A variation on this design involves the use of naturally occurring control groups rather than random assignment. Some programs, for example, have waiting lists or train groups in cycles. If strategies are used to ensure that the groups are comparable, the waiting list groups may serve as control groups for the groups who receive training. In fact, creating "waiting list" control groups (who will ultimately have the opportunity to receive training) rather than "no-treatment" controls may be a way to diminish ethical concerns about denying training to eligible individuals (Mark & Cook, 1984).

Conclusion

Longitudinal studies can provide valuable data about workplace literacy program outcomes. However, longitudinal designs also require more resources and commitment than

Figure 3

LONGITUDINAL EVALUATION DESIGN #3



3-15

short-term studies. They must be carefully designed and implemented; otherwise considerable time, energy, and resources will be spent to obtain information that is not very useful.

An additional strategy for enhancing the value of any longitudinal study is to gather qualitative as well as quantitative data. Qualitative data can be gathered through interviews and observations to provide greater information about the experiences of workers during and after training. It may offset the lack of control groups by offering insights into why certain outcomes do or do not occur.

Perhaps the key questions that must be addressed in a decision whether to use a longitudinal design are: What are the goals of the program? Are longitudinal methods needed to determine whether these goals are achieved? If so, then the results of longitudinal designs should justify the time and resources needed for a well-designed study.

References

Beder, H. (1991). *Adult literacy: Issues for policy and practice*. Malabar, FL: Krieger.

Hargroves, J. (1989, September/October). The basic skills crisis: One bank looks at its training investment. *New England Economic Review*, pp. 58-68.

Mark, M. & Cook, T. (1984). Design of randomized experiments and quasi-experiments. Chapter 4 in Rutman, L. (Ed.). *Evaluation research methods: A basic guide* (pp. 65-120). Beverly Hills, CA: Sage.

Menard, S. (1991). *Longitudinal research*. Newbury Park, CA: Sage.

Mikulecky, L. (1992). Models to evaluate effective workplace literacy programs. Chapter IV in Pelavin Associates, Inc. (Eds.). *Identifying effective adult education programs and practices*. (pp. IV-1 - IV-40). Washington, DC: U.S. Department of Education.

Pelavin Associates, Inc. (Eds.) (1992). *Identifying effective adult education programs and practices*. Washington, DC: U.S. Department of Education.

Young, M. (1992). Problems in conducting evaluations of adult education programs. Chapter VII in Pelavin Associates, Inc. (Eds.). *Identifying effective adult education programs and practices*. (pp. VII-1 - IV-13). Washington, DC: U.S. Department of Education.

ARTICULATION AND MEASUREMENT OF PARTICIPANT OUTCOMES IN WORKPLACE LITERACY PROGRAMS

Larry Mikulecky
Indiana University-Bloomington

Abstract

Much of the information presented in this paper is the result of a three-year study funded by the National Center on Adult Literacy to design, implement, and extend an evaluation model for workplace literacy programs (Mikulecky & Lloyd, 1992). For this NCAL project, the authors have designed a framework of interviews, questionnaires, productivity indicators, and guidelines for custom-designed workplace literacy assessments. Evaluation data from this model have now been collected from a variety of workplace literacy classes in six quite different workplace literacy programs (Mikulecky, 1992).

Foundation Elements of Workplace Literacy Programs to be Considered Before Evaluating Learner Impact

Two aspects of workplace literacy programs that directly influence program effectiveness and subsequent evaluations of program impact are the facts that:

(1) Nearly all workplace literacy programs involve negotiations among multiple stakeholders, and

(2) Typical classes are relatively brief in duration (i.e., 30-50 hours of instruction).

As background for the paper's discussion of participant outcomes, we provide a brief review of these foundation elements.

Multiple Stakeholders

Most workplace literacy programs are the results of joint efforts and partnerships. A survey of 107 workplace literacy program descriptions (Bussert, 1991) revealed that 92 percent were supported by two or more partners (i.e., businesses, unions, schools, and government agencies at local, state, and national levels). In addition, teachers (who are often hired part-time with little preparation time) are stakeholders since they are often left to decide what to teach and usually have their own ideas about what should be learned by "their learners."

Learners themselves must also be considered key stakeholders of workplace literacy programs. If the learner is unwilling to concentrate, invest mental energy, and practice outside of class, little can be accomplished. Within businesses, stakeholders often include both upper-level management (who want improved long-term employee performance) and line supervisors (who sometimes feel their jobs and short-term productivity are compromised by absent workers who are in classes). If training is believed to be linked to promotions and future job choice, organized labor also has a vested interest in the nature of training and recruitment.

Each of the above stakeholders can have different and potentially conflicting goals. For example, employers may desire improved productivity in current jobs or perhaps basic skills support for future training in such areas as quality assurance techniques, safety procedures, or use of new technologies. Some learners, in contrast, may desire training which

enables them to leave their current jobs or helps them achieve other personal goals. Funding agencies may have clear guidelines as to what sorts of instruction may or may not be offered. Gowen (1992) has documented some of the difficulties that can subvert program effectiveness when stakeholders are in disagreement about major program goals. Sometimes this can occur through direct conflict. At other times, committees can superficially agree that a brief instructional program should accomplish a dozen or more goals. It is often only when the program is in operation that stakeholders realize a brief program attempting to achieve a dozen goals is destined for failure. Either goals must be narrowed or short shrift is given to every goal, creating ineffective programs and angry stakeholders.

For the above reasons, an essential foundation element of any workplace literacy program evaluation must be determining if all key stakeholders have been involved with program planning. Extensive evaluation of a program's impact upon learners is not likely to be worth the effort if key stakeholders have not reached some consensus about a reasonable set of program goals which can be accomplished in the timeframes available. Indeed, funding programs for which key stakeholders haven't established a reasonable consensus about program goals is probably unwise.

What Can be Expected in Typical Timeframes

Literacy development takes time. An average child in school takes 180 days to improve a single grade level in literacy and computational abilities. Adults are able to learn more rapidly, but average adult basic education programs still take approximately 100 hours of practice for an adult to move from being able to read the relatively short sentences accessible to an average 4th grade reader to the slightly longer sentences accessible to an average 5th grade reader (Diekhoff, 1988; Mikulecky, 1989). Progressing from reading brief sentences to the point of being able to comprehend a plant newsletter, a memo, or a technical manual takes several hundred hours of practice. Documentation from the most effective adult and workplace literacy programs indicates they can help learners improve at nearly twice the rate of average programs (i.e., approximately 55 hours per grade level gain). In the very best of programs, it still takes a long time to move low-level literates to intermediate status and intermediate-level literates to the point of comprehending typical technical material.

Programs that focus tightly upon the vocabulary, formats, and specific sorts of literacy required for a limited type of reading (i.e., industry-specific tasks, safety information, etc.), can make slightly more rapid gains in targeted areas. There is only slight transfer, however, beyond the vocabulary and materials addressed by instruction.

This lack of easy transfer means that, unless programs can provide hundreds of hours of in-class and homework practice to improve broad-based general literacy, goals must be clear, focussed, and quite limited. Programs attempting to address too many or overly broad goals in a short time have virtually no likelihood of producing measurable learner impacts. Attempting to assess learner impact in such programs is likely to squander limited resources to no justifiable end.

Research on Effective Workplace Literacy Programs

Until recently, there was little research on effective workplace literacy programs. In a survey of the relatively few U.S. and Canadian workplace literacy programs for which there has been rigorous evaluation, Mikulecky and d'Adamo-Weinstein (1991:481) report:

(1) Effective programs require significant resources in terms of learner time on task (i.e. 50-100 hours of instruction per average 1 year of learner gain).

(2) Effective private programs report learner cost figures more than double those of average public programs (i.e., $7000 vs $2800 per learner).

(3) Effective programs integrate basic skills training with workplace technical training. This usually involves counseling as well as on-the-job training and analysis of the basic skills that learners need to perform their jobs.

This finding is supplemented by a study of 37 workplace literacy programs funded by the Department of Education (ED). The study identifies four key components of effective programs, including:

(1) Active involvement by project partners,

(2) Active involvement by employees in determining literacy needs,

(3) Systematic analysis of on-the-job literacy requirements, and

(4) Instructional materials directly related to the job (Kutner, Sherman, Webb, & Fisher, 1991).

No single class or course seems able to meet the demands of the diverse populations within a single workplace nor to provide a sufficient amount of instruction to move very

low-level literates to the functional literacy levels called for in developed nations' workplaces. Multi-strand approaches involving several different types of courses and strings of educational experiences leading to long-term training goals appear to offer the highest probability of success. In addition, the active involvement of workplace partners in developing programs and directly linking instruction to workplace demands appear to be key to program effectiveness. A relatively few workplace literacy programs meet all of these effectiveness criteria. Many ineffective programs simply have an outside provider deliver generic basic education in a workplace setting.

NCAL Workplace Literacy Impact Model

The National Center on Adult Literacy (NCAL) has funded the development of a model for providers of workplace literacy programs to evaluate the impact of these programs. This three-year project has developed techniques for gathering data on-site, piloted the model at two workplace programs, and demonstrated the feasibility of program providers using the model with minimal assistance at an additional four workplace programs

The impact assessment draws from a variety of research methodologies, including interview, observation, literacy task analysis, document analysis, survey research, traditional testing methodologies, and productivity monitoring techniques in current business use. The workplace literacy impact model focuses on learners (i.e., literacy improvement, changes in literacy practices, changes in beliefs about self and literacy, and changes in goals), upon employer objectives (i.e., improved safety, attendance, and productivity), and upon home and family literacy (i.e., changes in family literacy and educational practices).

The Conceptual Model

The conceptual model for the evaluation is based upon an expansion of Lytle's (1990) Beliefs, Practices, Processes, and Plans framework for categorizing changes in the relationships to literacy held by adults. Interviews, questionnaire items, and custom-designed job literacy tests assess learner changes in a variety of areas. An overview of those areas follows below.

Beliefs

- Perceptions of what is a literate individual
- Perceptions of self as literate
- Aspirations toward becoming more literate

Practices

- Amount and types of reading/writing at home
- Amount and types of reading/writing at work

Processes and Abilities

- Literacy strategies employed in think-aloud scenarios (i.e., When reading this newsletter or this graph, what would you do first, then next, then next?)
- Literacy comprehension of workplace materials using questions developed following a pattern for custom-designing workplace literacy tests
- Cloze test scores derived from workplace specific materials
- Standardized test, if general literacy is program goal

Plans

- Plans for self at 1, 5, and 10-year intervals
- Perceptions of literacy and education in those plans

Productivity

- Anchored supervisor ratings of job performance related to literacy and basic skills
- Attendance
- Safety
- Productivity suggestions made and approved
- Disciplinary measures and grievances

Family Literacy

- Literacy materials available in the home
- Literacy practices with one's children
- General literacy practices and modeling at home

The framework involves a shared mixture of interview and questionnaire items that are common across worksites as well as clear guidelines for developing parallel custom-designed assessments of learner literacy performance and productivity in using literacy and communication skills on the job. Shared items relate to the areas of beliefs, plans, literacy practices, and family literacy. Custom-designed items on workplace specific literacy improvement and productivity are developed using guidelines from the evaluation model to develop items parallel across worksites. This framework allows for program comparisons. As data from programs using the model accumulate, it is also possible for new program developers to establish more reasonable expectations of what instruction of a particular type and duration can achieve.

Rationale Behind Model Components

There is a two-fold rationale for assessing changes in learners' *beliefs and plans* related to literacy and learners' perceptions of their own literacy. Bandura's research on self-efficacy (1989) indicates that learners with higher perceptions of their own abilities (based upon accurate feedback) tend to try harder, continue in the face of obstacles, and succeed more often than learners of comparable tested ability but lower senses of personal effectiveness with literacy. Learners with low senses of personal effectiveness tend to subvert their own efforts with self-doubt and excuses for quitting. Low-literate adults tend to have exceptionally low and narrow visions of their own literacy abilities. Programs that help learners develop broader, more accurate senses of what it means to be literate and more detailed personal educational plans are more likely to have those learners use literacy outside of classrooms and continue learning after the completion of a class of brief duration. It is possible to quantify interview and questionnaire information in these areas and note the impact of an instructional program upon improved learner beliefs and educational aspirations.

In the areas of *literacy practices* and *family literacy*, questionnaire and interview items assess literacy activities over a specified period (usually seven days). It is possible to quantify results in terms of breadth, depth, and frequency of activities. Since literacy improvement takes a good deal of time, it is very important to determine if instruction increases literacy practice in the workplace, the home, and with one's family. Programs which foster literacy practice only in classrooms are severely limited in terms of expected learner gain. They simply cannot provide enough practice time to accomplish very much.

Assessing the impact of programs upon worker *productivity* has been very difficult--especially in workplaces which do not collect individual worker productivity data. In the relatively rare cases where such data are collected, it is possible to compare workers who have completed training to comparable workers waiting to take training or to workers' own previous performance. A broader definition of "productivity" can also assess the influence of workplace literacy programs upon employer goals such as improved attendance, better safety practices, increased participation in employee suggestion efforts, and so on. This information can also be compared to previous performance or to the performance of comparable workers if sufficient numbers of workers and long enough timeframes are available. Finally, supervisors or team leaders can work with program providers to develop pre and post anchored ratings of how well students enrolled in classes use literacy in workplace tasks compared to top, average, and below-average workers. Some examples of anchored-rating items developed by teachers and floor supervisors follow below.

Please rate each employee on a scale of 1 - 10 for each aspect below.

- An average employee would be rated 5.
- A top employee would be rated 8 or higher.
- A bottom employee would be rated 2 or lower.

COMMUNICATION

Bottom	**Average**	**Top**
won't speak; can't express self; nervous; won't shake hands	open, relaxed communicator; good listener and responder	processes information and responds with own analysis

1 2 3 4 5 6 7 8 9 10

PAPERWORK

Bottom	Average	Top
intimidated by job-related paperwork and does it poorly	does job-related paperwork, simply keeping pace procedures	completes all job-related paperwork and tries to improve

1 2 3 4 5 6 7 8 9 10

MACHINE SETTING

Bottom	Average	Top
unable to set machines correctly	usually sets machines correctly, but doesn't always check settings	sets machines correctly and checks settings thoroughly

1 2 3 4 5 6 7 8 9 10

Results of Evaluation Model Use

The evaluation model described above has now been used with a variety of class types at six quite different work sites selected for the wide range of programs. A listing of sites and types of classes at each site follows below.

1. Automotive Plant
   - Tech Prep (7 hours per day, 6 weeks)
   - GED (4 hours per week, 6 weeks)
   - ESL (8 hours per week, 6 weeks)
   - Control group for Tech Prep class

2. Wood Processing Plant
   - Communication & Collaboration (preparing self-directed teams, 4 hours per week, 8 weeks)

3. Women's Prison (Correctional Officers)
   - Writing Offender Reports (3 hours per week, 12 weeks)
   - Sgt.'s Exam Study Support (3 hours per week, 6 weeks)
   - Family Literacy (1 hour per week, 16 weeks)

4. Insurance Company
   - Job-related reading & writing (5 hours of class & home study, 8 weeks)

5. Hospital
   - Computer-based writing & study reading (2 hours per week, 10 weeks)

6. Gasket Maker
   - Basic reading & writing (4 hours per week, 16 weeks)

Results from analyzing pre and post data and, in some cases, comparing learner gains to those of a control group, indicate that instruction can produce improvement in all of the areas assessed on the evaluation model. No class and, indeed, no program demonstrates gains in all areas of the model. There is a pattern of gains being severely limited to areas directly addressed by instruction. Programs which spent time discussing the relevance of literacy and education tended to elicit changes in the areas of changed beliefs and aspirations. Programs that used workplace materials tended to have learners improve in the use of those materials and in literacy-related job performance. If instruction was limited to job-related prose literacy, improvement would occur with prose materials but not with charts or graphic materials. Improvement in family literacy practices occurred in the Women's Prison program, which had a family literacy component. It occurred slightly in the Insurance Company program, which required a significant amount of home study. Children tended to read more when parents were studying. There are a few exceptions to the pattern, but not many.

These observations about transfer need to be viewed somewhat cautiously. Data were collected primarily to test the workplace literacy impact model. No extensive observations of instruction were made. Class descriptions are based upon examination of curriculum materials and discussions with instructors. Still, ome suggestions for program expectations may be in order.

There seems to be little or no transfer of learning into areas not covered by instruction. For this reason, it appears that program providers need to have clear goals for what they want to achieve in the limited time (usually less than 100 hours) that learners are in class. They should also be seeking ways to extend this time beyond the classroom. One way of doing

this is to use on-the-job materials in class so that learners will be practicing outside class time. Also, encouraging learner motivation and independence is likely to lead to learners' engaging more often in literacy-related activities. The insurance program actually prepared take-home packets that doubled learning time and appeared to have some impact upon family literacy.

Use of the full evaluation model involves 20-30 minute interviews with learners before they enter programs and again after program completion. In addition, questionnaires and tests require an additional 20-30 minutes of learner time (usually possible in a group setting) before and after program involvement. If evaluation time needs to be curtailed, it seems reasonable to drop portions of the evaluation model unrelated to program goals and instruction.

Work with the evaluation model continues. A matrix of program data on electronic disk is being developed so new workplace programs can compare results to those of comparable programs that have already used the evaluation model. Further studies using and extending the evaluation model to establish the limits and parameters of transfer are also in the planning stages. In addition, a modified version of the evaluation model has been used with hundreds of learners in Adult Basic Education programs in 10 Indiana cities. Teachers in these programs have been trained to perform interviews, gather data, and write comprehensive evaluation reports using statistical results and a packet of guidelines and model reports.

References

Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology*, 25 (5) 729-35.

Bussert, K. (1991). *Survey of workplace literacy programs in the United States.* Unpublished manuscript. Language Education Department, Indiana University, Bloomington, IN.

Diekhoff, G. M. (1988). An appraisal of adult literacy programs: Reading between the lines. *Journal of Reading, 31*, 624-630.

Gowen, Sheryl G. (1992), *The politics of workplace literacy: A case study.* New York: Teachers College, Columbia University.

Kutner, Mark, Sherman, Rene, Webb, Lenore, & Fisher, Caricia. (1991). *A review of the national workplace literacy program.* Washington, DC: U.S. Department of Education, Office of Planning, Budget and Evaluation.

Lytle, S. L. (1990). *Living Literacy: The practices and beliefs of adult learners.* Presented at an Invited Symposium of the Language Development SIG entitled: "Adult Literacy/Child Literacy: One World or Worlds Apart." American Educational Research Association Annual Meeting, Boston, MA.

Mikulecky, L. J. & D'Adamo-Weinstein, L. (1991). How effective are workplace literacy programs? In Taylor, M., Lewe, G., and Draper, J. (Eds.) *Basic Skills for the Workplace*. Toronto: Culture Concepts, Inc.

Mikulecky, L. J. & Lloyd, D. P. (1992). *Evaluating the Impact of Workplace Literacy Programs: Results and Instruments from the NCAL Workplace Literacy Impact Project.* Bloomington, In: Indiana University School of Education.

Mikulecky, L. J. (1989). Second chance basic skills education. In *Investing in People, Commission on Workforce Quality and Labor Force Efficiency*, Washington, DC: U.S. Department of Labor, Volume 1, pp. 215-258.

Mikulecky, L. (1992). Workplace literacy programs: Variations of approach and limits of impact. A paper presented at the National Reading Conference annual meeting, San Antonio, TX. December 3, 1992.

ARTICULATION AND MEASUREMENT OF PROGRAM OUTCOMES

Anthony R. Sarmiento
Education Department, AFL-CIO

Abstract

The National Workplace Literacy Program (NWLP) is designed to improve workforce productivity through improving workers' literacy skills. However, after five years of NWLP demonstration projects, the available evaluation literature is limited, and relatively little knowledge has been gained about the impacts of these programs on participating employers and employees. This paper employs available research on general training and education programs for incumbent workers to describe program outcomes for evaluations of the NWLP. Related federal programs, such as those funded under the Job Training Partnership Act (JTPA), typically target different populations (and thus outcomes) from those of programs funded under NWLP, emphasizing job placement and earnings gains for *unemployed* workers.

State efforts, on the other hand, include many training programs in job-related basic skills (including literacy) for existing employees. Studies of these programs highlight the important substantive and methodological factors in evaluating government-funded training programs, including measuring program outcomes such as training objectives, company and unit performance, and earnings. However, the studies also highlight the rudimentary knowledge available on workplace training, including limited agreement over its objectives and how to measure their accomplishment, especially when constructs such as workplace orientation, structure, and productivity are considered. The paper concludes with several implications for National Workplace Literacy Program administrators and evaluators, including recommendations that NWLP training objectives be articulated more clearly and that federal program administration be predicated on a broader interpretation of the statutory purposes of the legislation that can be actively embraced by policymakers, educators, employers, and unions.

Introduction

It is clearly in the public interest to invest more effort and resources into evaluating the U.S. Department of Education's National Workplace Literacy Program (NWLP). After five years of demonstration projects, it is not readily apparent what has been learned about the impact of workplace literacy programs on either participating employers or participating learners. The existing program evaluations rely primarily on self-report data rather than empirical or quantitative research. Since the NWLP represents the federal government's largest effort to improve the skills and productivity of incumbent workers, a more rigorous evaluation of its projects would be of unquestionable value.

But it is important to keep in mind that several methodological and substantive problems lie ahead in our effort to design an evaluation that will meet the needs of literacy policymakers as well as practitioners. This situation is made more complicated by the separation which has been made between program and participant outcomes. This issues paper aims to identify some of these key problems as indicated by a review of the literature on training and education program evaluation, productivity, and workplace restructuring. The paper concludes with several implications for the evaluators and administrators of NWLP.

Summary of Available Research

Because the available literature on workplace literacy program evaluations is so limited, this paper expands the scope of inquiry, drawing from research conducted on similar training and education programs. Rather than look at evaluations of other types of literacy programs, it was more useful to define workplace literacy programs as a subset of the larger set of programs that train incumbent workers. Specifically, the evaluation literature on the following types of programs for incumbent workers was examined:

- Privately supported training
- State government financed training.

Perhaps not surprisingly, this survey found that workplace literacy program evaluation reflects many of the same trends and issues as this larger body of evaluation research. Namely, there is very little quantitative evaluation of either publicly or privately funded

training programs, especially impact on employer or workplace objectives. Furthermore, there did not appear to be a clear trend or strong support for expanding this kind of program outcome evaluation.

Evaluations of Privately Supported Training for Incumbent Workers

The evaluation literature describing training programs funded by private employers and unions is dominated by a framework proposed by Kirkpatrick in the late 1960s. His four-tiered model remains the most widely known and accepted framework for evaluating incumbent worker training programs. Ideally, an evaluation should measure:

Reaction: How well did training participants like the program?

Learning: What knowledge (principles, facts, and techniques) did participants gain from the program?

Behavior: What positive changes in participants' job behaviors stemmed from the training program?

Results: What were the training program's organizational effects in terms of reduced costs, improved quality of work, increased quantity of work, and so forth (Carnevale and Schulz, 1990)?

Kirkpatrick's hierarchical model continues to be promoted by major associations of training and personnel administration professionals. The American Society for Training and Development (ASTD) has described the Kirkpatrick model as "the evaluation framework that most training practitioners use" (Carnevale and Schulz, 1990). The Society for Human Resource Management (SHRM, formerly the American Society for Personnel Administration) describes it as "the most widely accepted framework used by training researchers in evaluating training programs" (Grove and Ostroff, 1991). Recent articles in an independent publication, *Training*, also frequently cite the Kirkpatrick model (Gordon, 1991, and Hassett, 1992).

Although there seems to be a clear consensus among employers on what constitutes good training evaluation design, most private sector employers rely on subjective data to assess their own training programs--if they evaluate at all. Relatively few evaluations attempt to gather quantifiable measures or follow rigorous experimental designs.

In 1978, Kirkpatrick surveyed more than 100 firms that evaluated their training programs and found that these evaluations contained the following:

Measurement of participant reactions	75 percent
Measurement of learning	<50 percent
Measurement of behavioral changes	<20 percent
Measurement of results on the job	15 percent

Two studies conducted about 10 years later found similar patterns. As cited by the Office of Technology Assessment study on worker training, Meigs-Burkhart reported in 1986 that a survey of training professionals in major corporations indicated that less than half of their training programs were evaluated at all. A year later, a survey of 43 companies found that most respondents used the following evaluation criteria (Clegg, 1987):

- Informal collection of passing comments
- End of course evaluation sheets
- Instructor reports of training success
- Continued demand for courses.

In 1988, ASTD conducted a poll of organizations that led in training evaluation. They found that only 20 percent studied the economic effects of training (Carnevale and Schulz, 1990).

Both SHRM and ASTD maintain that determining the return on training investments is becoming increasingly important, but they also have found that employers do not appear to support macro-level evaluations which examine training results (Kirkpatrick's fourth tier). SHRM concluded in 1991 that "specific measures of and research on training effects at the macro level are fairly scarce" (Grove and Ostroff, 1991), while ASTD's Carnevale and Schulz arrived at a similar conclusion:

> Despite the growing demand for accountability, financial accounting for training shows only a slight increase. . . .when it comes to investments in training and development, subjective decisions prevail" (Carnevale and Schulz, 1990).

Evaluations of Publicly Supported Training for Incumbent Workers

As noted earlier, the NWLP is the largest federally funded program designed to train and educate incumbent workers. Unlike much larger programs like the Job Training

Partnership Act (JTPA) or Even Start, the NWLP is not income-targeted or limited to serve only those adults who are economically disadvantaged or unemployed.

As a result, the evaluation literature describing programs under JTPA (and in the past, CETA) and similar federal initiatives provides only limited guidance in designing NWLP evaluations. A key difference is that the most important program outcomes for training supported under JTPA are not analogous to NWLP. Most federal training programs emphasize job placement and/or earnings gains, while most, if not all, participants in NWLP training programs are already employed.

Although the NWLP is unique among federal education and training programs, state governments are funding more training programs aimed at incumbent workers. The most recent studies estimate that since the early 1980s 46 states have established one or more customized training programs for existing employees. Training in job-related basic skills is usually an allowable activity in such programs (Creticos and Sheets, 1990).

While the programs vary in size, emphasis, organizational structure, and experience, they represent a significant aggregate effort by government. The larger individual state programs are especially noteworthy. For instance, the California Employment and Training Panel spent $300 million serving over 160,000 trainees between 1983 and 1991. The Industrial Training Program in Illinois spent $106 million between 1979 and 1990 to train 130,000 workers in 700 firms (Batt and Osterman, 1993b).

But these state-financed efforts follow the same research pattern as their private-sector counterparts: formal program evaluations are equally scarce. While there is growing business interest in such programs and calls to increase funding levels, critics argue that there is insufficient evidence of the programs' effectiveness, with some suggesting that public funds are merely being substituted for private training funds.

As state agencies responsible for these programs were increasingly asked to demonstrate their effectiveness, a joint study was initiated in 1989 by the National Governors' Association (NGA) with support from the National Commission for Employment Policy (NCEP). The study had three objectives:

- To clarify the policy rationale and intervention model for state-financed, workplace-based retraining programs in terms of their twin objectives of retaining jobs and preventing unemployment;

- To develop an evaluation system that would include:
  - business screening guidelines for targeting training investment and minimizing substitution effects, and
  - a performance outcome system for measuring program outcomes that are necessary to achieve the twin objectives, such as training (behavioral) objectives, work unit performance, company performance, and trainee earnings; and
- To examine the feasibility of implementing the proposed evaluation system in four different state programs.

The study was completed and issued its final report in 1992. The findings and recommendations represent a significant advance in identifying the substantive and methodological issues in evaluating government-funded training programs for incumbent workers. But the reports also reveal that the emerging field is starting from a fundamental level of inquiry. Also, it is too early to tell how widely the recommendations will be embraced and implemented by the various state programs.

Evaluating Training Program Outcomes: An Emerging Field

In 1990 researchers reported to the Industrial Relations Research Association (IRRA) the following observation:

> Considering the massive volume of training ongoing in the U.S. economy and the constant advocacy for more, it may be surprising how little is actually known about the results. . . .
>
> Training and development experts in private industry are only beginning to discuss the efficacy and the methodology for outcome evaluations to determine the rate of return. Those within public-sector employment, including the military, seem even less likely to raise the issue (Mangum et al., 1990).

This conclusion was reached after the researchers reviewed the literature on rates of return to in-school and post-school occupational training, military training, publicly funded training for economically disadvantaged and displaced workers, and employer-sponsored training. The researchers found that evaluations which examined the rate of return were largely limited to publicly supported vocational education and training programs for

economically disadvantaged persons. For all other programs, training appears to be "an act of faith."

These researchers identified at least four major reasons for the paucity of empirical research:

1. A lack of agreement over training objectives;
2. Interdisciplinary conflicts over what to measure and how to measure it;
3. Technical difficulties in measuring and isolating the effects of training; and
4. Inadequate support among decisionmakers to obtain sophisticated rate of return information about training programs.

There have been noteworthy efforts to address some of these obstacles. For instance, with funding from the U.S. Department of Labor, in 1990 ASTD developed a consensus accounting model with the help of training and accounting experts who recognized that "there were no generally accepted accounting procedures for valuation of human assets." It is premature to judge whether their recommendations will become the norm.

Other obstacles still remain--particularly the conflict between the technical requirements of experimental or even quasi-experimental designs and the realities of actual workplaces. Many employers and workers would not tolerate measures required to collect quantifiable data or minimize selectivity bias. In addition, it would be nearly impossible to control for numerous external factors: the changes in supervision, management, competitive strategies, customer demands, and state of the economy which affect workplaces today.

While the NGA study of state-financed retraining programs found that net impact evaluations were rare, it concluded that "impact evaluation systems based on either experimental or quasi-experimental designs are not feasible at this time (Creticos & Sheets, 1990)."

Worker Productivity and Training

Although the relationship between literacy skills and productivity may appear simple and direct, there is a growing consensus that improving workers' occupational or literacy skills may be a necessary, but not a sufficient, requirement for improving productivity.

The NGA study provides an apt description of what is involved:

> . . .worker productivity is a function of both worker skills and the integration of these skills into a productive workplace that incorporates appropriate process technologies, job design, labor-management relations, compensation systems, and employment security policies. To be successful, employee retraining must be integrated with larger workplace changes that are intended to improve work unit or company performance (Creticos, 1990).

Other reports like *America's Choice: High Skills or Low Wages!* have made similar arguments that the nation needs to restructure workplaces as well as prepare its workers better. More literate or better skilled workers will have little impact on productivity unless more workplaces are transformed from the traditional, mass production model to "high performance" companies. The estimates of how many workplaces have been transformed so far range from five to 30 percent.

But even if one accepts the most optimistic estimates by Paul Osterman and others (forthcoming), one is still left with the reality that two of every three workplaces in the United States today are structured on a management philosophy that minimizes the importance of worker skills. Governed by the pressures to achieve immediate bottom-line returns, training is provided only if it cannot be avoided, and the shorter the training the better.

Recently in ASTD's Training and Development Journal, two training consultants provided the following advice in an article, "Demonstrating ROI of Training:"

> The easiest ways to keep training costs low are to reduce trainee and instructor hours. . . .Another effective way to reduce costs is to simplify, automate, re-engineer, or eliminate job tasks. It may be more cost-effective to eliminate resource and procedural problems than to train employees to work around them.

Limited research has been conducted to determine whether training programs lead to workplace transformation, or the reverse. Batt and Osterman have argued that there is no reason to give skill any particular causal primacy. In fact, several case studies of American firms that have restructured clearly indicate that the firm committed itself to transformation first, then decided to increase its training investments.

Additional case studies have concluded that training prior to transformation does little for workers or their employers. Foucar-Szocki's report on a field test of ASTD's Workplace Basics model in two Virginia firms advised that "enhancing skills in an environment where the newly-developed skills cannot be used is counterproductive."

Implications for the National Workplace Literacy Program

By law, the purpose of the National Workplace Literacy Program is "to improve the productivity of the workforce through improvement of literacy skills needed in the workplace." After reviewing the literature on evaluating programs with similar purposes--as well as some of the recent debates about productivity and workplace transformation--one can better understand why it will not be easy to design an evaluation of the NWLP that meets the needs of literacy policymakers and practitioners.

It is helpful to recall the four obstacles identified by the IRRA that have discouraged rate-of-return evaluations of most training and education programs. While the technical difficulties related to research methodology cannot be avoided indefinitely, other issues that are more political in nature should be addressed at this stage in the history of NWLP. Namely, there needs to be a stronger and clearer articulation of the training objectives of the NWLP which is endorsed by all the key players: educators, employers, unions, and the federal government. Without this consensus or shared vision, there will be little support or will to overcome the other obstacles that prevent rigorous and useful evaluations of the NWLP.

In particular, the U.S. Department of Education should reevaluate its interpretation of the statutory purpose of the NWLP. The regulations and rules governing the program reflect a narrow or strict interpretation of the law that established NWLP. This trend is evident as the Invitations for Proposals have changed over the program cycles and in the most recent amendments to the regulations proposed in May 1993. Instead of supporting workplace programs that develop broad literacy skills through a variety of activities, the Department has focused almost exclusively on those literacy skills needed by the actual jobs held by the participating learners.

The authorizing legislation specifies six types of allowable activities, of which teaching job-specific literacy skills is only one:

A. Providing adult literacy and other basic skills services and activities;

B. Providing adult secondary education services and activities which may lead to the completion of a high school diploma or equivalent;

C. Meeting the literacy needs of adults with limited English proficiency;

D. Upgrading or updating basic skills of adult workers in accordance with changes in workplace requirements, technology, products, or processes;

E. Improving the competency of adult workers in speaking, listening, reasoning, and problem solving, **or** (emphasis added)

F. Providing education counseling, transportation, and nonworking hours child care services to adult workers while they participate in a program.

Given what is known about productivity and worker training, the U.S. Department of Education has firm grounds on which to administer the National Workplace Literacy Program based on a broader interpretation of its statutory purpose. By rejecting a more narrow view, the Department would also establish more room to build a consensus on the NWLP's training objectives which could be actively embraced by educators, employers, and unions. Taking this step would enable the workplace literacy field to overcome one of the obstacles to better program evaluations, and move the nation a little closer to establishing a coherent and comprehensive system of lifelong learning.

References

Batt, R., & Osterman, P. (1993a). *A national policy for workplace training: Lessons from state and local experiments*. Washington, DC: Economic Policy Institute.

Batt, R., & Osterman, P. (1993b). *Workplace training policy: Case studies of state and local experiments* (Working paper no. 105) Washington, DC: Economic Policy Institute.

Carnevale, A. P., & Schulz, E. R. (1990). Technical training in America: How much and who gets it? *Training and Development Journal, 42*, 18-32.

Carnevale, A. P., & Schulz, E. R. (1990). Return on investment: Accounting for training. *Training & Development Journal (Supplement)*, 41, S-1 - S-32.

Clegg, W.H. (1987). Management training evaluation: An update. *Training and Development Journal; 41*, 65-71.

Commission on the Skills of the American Workforce. (1990). *America's choice: High skills or low wages!* Rochester, NY: National Center on Education and the Economy.

Creticos, P.A., & Sheets, R.G. (1990). *Evaluating state-financed, workplace-based retraining programs: A report on the feasibility of a business screening and performance outcome evaluation system.* (Research Report 89-08). Washington, DC: National Commission for Employment Policy.

Creticos, P. A., & Sheets, R. G. (1992). *Evaluating state-financed, workplace-based retraining programs: Case studies of retraining projects.* (Research Report 91-05) Washington, DC: National Commission for Employment Policy.

Davidove, E. A., & Schroeder, P. A. (1992). Demonstrating ROI of Training. *Training and Development Journal, 46*(8), 70-71.

Foucar-Szocki, D. L. (1992). *Beyond training: A field test of the American Society for Training & Development's workplace basics*. Staunton, VA: Education & Training Corporation.

Gordon, J. (1991). Measuring the 'goodness' of training. *Training, 28*, 19-26.

Grove, D. A., & Ostrofi, C. (1991). Program Evaluation. In K. Wexley & J. Hinrichs (Eds.), *Developing Human Resources* (pp. 5-158 - 5-220). Washington, DC: Bureau of National Affairs.

Hassett, J. (1992). Simplifying ROI. *Training, 29*, 53-58.

Kirkpatrick, D. L. (1978). Techniques for evaluating training programs. *Training and Development Journal, 33(6),* 78+.

Mangum, S., Mangum, G., & Hansen, G. (1990). Assessing the Returns to Training. In L. Ferman, M. Hoyman, J. Cutcher-Gershenfeld, & E. J. Savoie (Eds.), *New Developments in Worker Training: A Legacy for the 1990s* (pp. 55-89). Madison, WI: Industrial Relations Research Association.

Meigs-Burkhart, T. (1986). *Employee training in America: A comparative assessment of training and development.* Princeton, NJ: Opinion Research Corporation.

Osterman, P. (In Press). How Common is Workplace Transformation and How Can We Explain Who Does It? Results From a National Survey. *Industrial and Labor Relations Review.*

U.S. Congress, Office of Technology Assessment. (1990). *Worker training: Competing in the new international economy.* (OTA-ITE-457). Washington, DC: U.S. Government Printing Office.

Appendix A

Work Group Conference Agenda

DESIGN GUIDANCE FOR EVALUATING WORKPLACE LITERACY PROGRAMS

Work Group Conference Agenda

Dupont Plaza Hotel
1500 New Hampshire Avenue, NW
Washington, DC 20036

April 13, 1993

8:30 - 9:00 Continental Breakfast

9:00 - 9:15 Introduction and Welcome

Sandra Furey
Planning and Evaluation Service, Office of Policy and Planning,
U.S. Department of Education

Ronald S. Pugsley
Division of Adult Education and Literacy, Office of Vocational and Adult Education, U.S. Department of Education

9:15 - 10:30 *Key Components of Workplace Literacy Projects and Definition of Project "Models"*

Judy Alamprese, COSMOS Corporation

10:30 - 10:45 Break

10:45 - 12:00 *Evaluation Purposes and Methodological Alternatives*

Jorie Philippi, Performance Plus Learning Consultants

12:00 - 1:15 Lunch

1:15 - 2:30 *Advantages and Disadvantages of Longitudinal Designs for Evaluating Workplace Literacy Programs*

Elisabeth Hayes, University of Wisconsin, Madison

Work Group Conference Agenda
Page 2

2:30 - 3:15 *Articulation and Measurement of Program Outcomes*

Tony Sarmiento, AFL-CIO

3:15 - 3:30 Break

3:30 - 4:45 *Articulation and Measurement of Participant Outcomes*

Larry Mikulecky, University of Indiana

4:45 - 5:15 *Summary and Synthesis of Key Points*

Thomas Sticht, Applied Behavioral and Cognitive Science

Appendix B

Biographies of Presenters

Biographies of Presenters

Judith A. Alamprese
COSMOS Corporation

Judith Alamprese is Director of the Education and Training Group at the COSMOS Corporation, where she is responsible for initiatives in education, policy analysis, and job training research. She has directed over 20 projects in the fields of adult literacy and basic skills, job training, partnerships, professional development, and student assessment. She has also served as the Chair of the Research Task Force of the National Institute for Literacy, and as advisor to the U.S. Department of Labor's Workplace Literacy Project and the Kettering Foundation's National Literacy Issues Forum. She is currently assessing evaluation methodologies used by National Workplace Literacy Program grantees for the U.S. Department of Education.

Ms. Alamprese has developed a number of research and evaluation reports in the field of adult basic skills and literacy education. A sampling includes *Assessing the Effects of Volunteer Literacy Programs: Current Practice and Future Directions; Study of Federal Funding Sources and Services for Adult Education; Direct and Equitable Access: Collaborative Opportunities Under the National Literacy Act;* and *Partnerships for Change: Strategies for Improving Workplace Literacy and Agricultural Upgrade Opportunities for Migrant and Seasonal Farmworkers.*

Elisabeth R. Hayes
University of Wisconsin-Madison

Elisabeth R. Hayes is Assistant Professor in the Department of Continuing and Vocational Education at the University of Wisconsin-Madison, where her research focuses on gender and racial issues in adult education. Prior to her faculty appointment, Dr. Hayes served as project director, assessment specialist, and basic skills instructor in adult literacy programs.

Dr. Hayes has designed several basic skills and literacy programs. For example, she developed U.S. Department of Education-funded Student Literacy Corps Programs at two universities, and routinely offers Adult Basic Education Institutes. She has developed courses in "Methods of Adult Literacy Education," "the Adult Learner," "Principles and Practices of Adult Basic Education," and "Questionnaire Development."

Dr. Hayes has published numerous articles, chapters, and reports on adult and continuing education, including literacy and basic skills education. These include *Literacy and Women's Lives: Old Problems, New Possibilities; Functional Literacy Needs of Low Literate Adult Basic Education Students; Ends and Means of Adult Literacy Education; Typology of Low-literate Adults Based on Perceptions of Deterrents to Participation in Adult Basic Education; Evaluation of the Effectiveness of Basic Skills Programs with an Academic Focus vs. Functional/Competency-Based Programs;* and *A Brief Guide to Critiquing Research.*

Larry Mikulecky

Indiana University

Larry Mikulecky is Professor of Education and Chair of the Language Education Department at Indiana University, where his research examines the literacy requirements for success in the military, business, universities, and secondary schools. He has served as principal investigator on numerous research projects funded by the U.S. Departments of Education and Labor, foundations, and corporations. He is currently directing the National Center on Adult Literacy's Workplace Literacy Impact Evaluation studies.

Dr. Mikulecky has served as an international training, evaluation, and document design consultant in Australia and Canada. He has worked on or directed projects for the U.S. military, U.S. Departments of Education and Labor, the Ontario Ministry of Skills Development, United Auto Workers/Ford, Business Council for Effective Literacy, and numerous corporations and school districts.

Dr. Mikulecky has published over 100 journal articles, chapters, and texts. For example, he is lead author on the recent Simon and Schuster series on *Strategic Skill Builders for Banking*, as well as Cambridge Publications basic skills series *On the Job*.

Jorie Wilkinson Philippi

Performance Plus Learning Consultants

Jorie W. Philippi is the founder and Executive Director of Performance Plus Learning Consultants, Inc. She has extensive experience in basic skills education and remediation, curriculum development, training, and evaluation. She served as director of the basic skills remedial reading program for U.S. Army enlisted personnel in Europe, where she developed job-specific curriculum and standardized tests, led teacher inservice training on effective instructional strategies for adults, and evaluated the program.

Ms. Philippi designed numerous functional context basic skills curricula for educationally disadvantaged young adults. She currently chairs the Adult Literacy Committee of the International Reading Association and the Business and Industry Unit of the American Association of Adult and Continuing Education. She is also a member of the National Advisory Committee for the Work in America Institute's Study of Job-Linked Literacy Programs.

Ms. Philippi consults on workplace literacy issues with businesses, industries, labor, and educational and professional organizations. She has written a number of texts and articles, including a workplace literacy "how-to" manual, *Literacy at Work: The Workbook for Program Developers*, and a video-based teacher training course, *Retraining the Workforce: Meeting the Global Challenge*.

Anthony R. Sarmiento
AFL-CIO

Anthony R. Sarmiento has been assistant director of the Education Department of the AFL-CIO since 1990. He currently represents the AFL-CIO on the Literacy Definition Committee for the U.S. Department of Education's National Adult Literacy Survey, the National Advisory Panel of the National Center on Adult Literacy, and the GED Testing Service Advisory Committee.

Before joining the AFL-CIO Education Department, Mr. Sarmiento served 11 years with the AFL-CIO Human Resources Development Institute, where he directed national research and demonstration projects in workplace literacy, career-ladder training programs, school-to-work transition programs, and displaced worker assistance and training centers. Prior to joining the AFL-CIO, Mr. Sarmiento worked for the District of Columbia government, where he developed a number of innovative education and training programs for economically disadvantaged youth that involved partnerships with the private sector.

Mr. Sarmiento has published extensively on the implications of education policy for the labor movement, including *Worker-Centered Learning: A Union Guide to Workplace Literacy; Do Workplace Literacy Programs Promote High Skills or Low Wages? Suggestions for Future Evaluations of Workplace Literacy Programs; Wanted: A New Kind of Workplace;* and *Workplace Literacy, Workplace Politics.*

Thomas Sticht
Applied Behavioral and Cognitive Sciences

Thomas Sticht is President and Senior Scientist at Applied Behavioral and Cognitive Sciences, Inc. His research and development efforts have focused on the application of cognitive science and communications technology to the education and training needs of underserved youth and adults.

Prior to his current position, Dr. Sticht was Vice President of the Basic Skills Division of the Human Resources Research Organization. He also served as Associate Director of the National Institute on Education, where he directed the research and development programs of the Basic Skills Section. He has served on the faculty at the Harvard University School of Education. Dr. Sticht currently serves as the chief technical adviser to the Work in America Institute's Job-Linked Literacy project; he is a member of the selection jury for UNESCO Literacy Prizes, the National Governors' Association's Advisory Group for Achieving National Education Goal 5 on Adult Literacy, and the National Advisory Board of the Business Council for Effective Literacy.

He has published over a hundred reports, articles, and books, including *Evaluating National Workplace Literacy Programs; Reading for Working: A Functional Literacy Anthology; Literacy and Human Resources Development at Work;* and *Testing and Assessment in Adult Basic Education and English as a Second Language Programs.*

Appendix C

Work Group Participants

DESIGN GUIDANCE FOR EVALUATING WORKPLACE LITERACY PROGRAMS

Work Group Participants

Nancy Adelman Policy Studies Associates 1718 Connecticut Avenue, NW Washington, DC 20009	Marty Cohen Work In America Institute 700 White Plains Road Scarsdale, NY 10583
Judy Alamprese COSMOS Education Training Group 1735 I Street, NW, #613 Washington, DC 20006	Thomas R. Curtin Research Triangle Institute P.O. Box 12194 Research Triangle Park, NC 27514
Bob Bickerton Massachusetts State Dept. of Education Quincy Center Plaza 1385 Hancock Street Quincy, MA 02169	Barbara Elliott Research Triangle Institute P.O. Box 12194 Research Triangle Park, NC 27514
Miriam Burt Food and Beverage Workers Union, Local 32 1221 Massachusetts Avenue, NW Ground Level Washington, DC 20005	Sandra Furey Office of Planning and Policy U.S. Department of Education 400 Maryland Avenue, SW Washington, DC 20202
Donna Cheatham Pennsylvania Blue Shield Continuing Education 1800 Center Street Camp Hill, PA 17089	Elisabeth Hayes University of Wisconsin Dept of Continuing and Vocational Education 225 N. Mills Street Madison, WI 53706
Forrest Chisman Southport Institute for Policy Analysis 820 First Street, NE, Suite 468 Washington, DC 20002	Becky Hayward Research Triangle Institute P.O. Box 12194 Research Triangle Park, NC 27514
Carol Clymer-Spradling El Paso Community College P.O. Box 20500 El Paso, TX 79998	Paul Jurmo Literacy Partnerships 21 Van Houten Avenue Jersey City, NJ 07305

Work Group Participants (continued)

Mark Kutner Pelavin Associates 2030 M Street, NW, Suite 800 Washington, DC 20036	Delores Perin Center for Advanced Study in Education CUNY Graduate School, Room 620 25 W 43rd Street New York, NY 10036
Inaam Mansoor Arlington Education & Employment Program Wilson School 1601 Wilson Boulevard Arlington, VA 22209	Jorie Philippi Performance Plus Learning Consultants 6 California Avenue Charleston, WV 25311
Larry Mikulecky School of Education Room 3038 Indiana University Bloomington, IN 47405	Ron Pugsley Office of Vocational and Adult Education U.S. Department of Education 400 Maryland Avenue, S.W. Washington, DC 20202
Mary Moore Mathmatica Policy Research 600 Maryland Avenue, SW, Suite 550 Washington, DC 20024	Nancy Rhett Office of Planning and Policy U.S. Department of Education 400 Maryland Avenue, S.W. Washington, DC 20202
Melissa Morrill Office of Vocational and Adult Education U.S. Department of Education 400 Maryland Avenue, SW Washington, DC 20202	Tony Sarmiento AFL-CIO Education Department 815 16th Street, NW Washington, DC 20006
Sarah Newcomb Office of Vocational and Adult Education U.S. Department of Education 400 Maryland Avenue, SW Washington, DC 20202	Joel Sherman Pelavin Associates 2030 M Street, NW, Suite 800 Washington, DC 20036
Audrey Pendleton Office of Planning and Policy U.S. Department of Education 400 Maryland Avenue, SW, Room 3141 Washington, D.C. 20202	Laura Sperazi Evaluation Research 130 Warren Street Newton Center, MA 02159

DESIGN GUIDANCE FOR EVALUATING WORKPLACE LITERACY PROGRAMS

Work Group Participants (continued)

Tom Sticht Applied Behavioral and Cognitive Sciences 2062 Valley View Boulevard El Cajon, CA 92019	Carol Young Civil Service Employees Association LEAP 143 Washington Avenue Albany, NY 12210
Ricky Takai Acting Assistant Secretary Office of Vocational and Adult Education U.S. Department of Education 330 C Street, SW Washington, DC 20202	Malcolm Young Development Associates 1730 N. Lynn Street Arlington, VA 22209
Pamela Wall ABCs of Construction Project 3028 Bell Cherie Baton Rouge, LA 70820	David Zacchei The Travelers 1 Tower Square, 5 TEC Hartford, CT 06183